LEVEL

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD CONFERENCE PROCEEDINGS No. 230
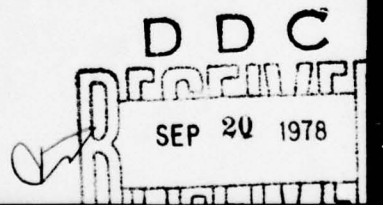
Impact of Charge Coupled Devices and Surface Acoustic Wave Devices on Signal Processing and Imagery in Advanced Systems

NORTH ATLANTIC TREATY ORGANIZATION

**DISTRIBUTION AND AVAILABILITY
ON BACK COVER**

78 09 18 196

AGARD-CP-230

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

AGARD Conference Proceedings No.230

IMPACT OF CHARGE COUPLED DEVICES AND SURFACE ACOUSTIC WAVE DEVICES ON SIGNAL PROCESSING AND IMAGERY IN ADVANCED SYSTEMS

Edited by

Y. Brault
Thomson CSF
Division Equipements Avioniques & Spatiaux
178 Bld Gabriel Péri
92240 Malakoff

15 oct 77   490 p.

ACCESSION for

NTIS   White Section ☑
DDC   Buff Section ☐
UNANNOUNCED   ☐
JUSTIFICATION

BY
DISTRIBUTION/AVAILABILITY CODES

DIST   AVAIL. and/or SPECIAL

A

Papers presented at the Avionic Panel Symposium held in Ottawa, Canada, 11–15 October 1977.

400 043 78 09 18 196

THE MISSION OF AGARD

The mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

— Exchanging of scientific and technical information;

— Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;

— Improving the co-operation among member nations in aerospace research and development;

— Providing scientific and technical advice and assistance to the North Atlantic Military Committee in the field of aerospace research and development;

— Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field;

— Providing assistance to member nations for the purpose of increasing their scientific and technical potential;

— Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Program and the Aerospace Applications Studies Program. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

PREFACE

The general goals of the AGARD Avionics Panel in organizing this symposium will be better understood by recalling first the guidelines given to the authors.

- Papers were requested to be mainly devoted to practical considerations, applications, and results.
- The impact of the new techniques, CCD and SAW, on the design of future systems had to be emphasized.
- Specific efforts were required in respect of:
  - contrasting SAW and CCD and defining the areas of applications for each.
  - contrasting these new techniques with the overwhelming digital technique and indicating whether these techniques are competing or are complementary.
  - describing near-term improvements and developments.
  - explaining the advantages (in terms of size, weight, performance, and cost) expected from their eventual extensive use.

Answers to the above questions are of vital importance for system designers who are faced today with extremely fast improvements in available technologies. A peremptory need exists to provide them with a clear understanding of what can be offered to save weight, size and costs in avionics systems, as these parameters are more and more important in airborne equipments.

It can be stated that the goals have been reached during this symposium when considering the high quality of papers submitted, and the high attendance of electronic engineers and scientists coming from all the NATO countries. The time devoted to discussions following the papers was somewhat too short.

The symposium was divided into five sessions:

Session 1 — Fundamental Technology and Trends

Session 2 — Technology and Devices

Session 3 — Fundamental modules/subsystems

Session 4 — Imagery

Session 5 — Signal processing.

Session 1 — The symposium was opened by a noteworthy review of the state-of-the-art of CCD and SAW technologies, and their roles in signal processing and imagery. Two invited papers were presented; one by Lewis T. Claiborne (U.S.A.) and the second by Dr. J.B.G. Roberts (U.K.).

- The two authors provided converging views on the coexistence of CCD, SAW and digital techniques as each of them has its inherent advantage without precluding the other's use. However, it must be noted that if SAW are far more advantageous than digital techniques in specific applications, it is not so clear for CCD technologies, the choice between CCD and digital techniques needing a careful case-by-case analysis.

- Digital techniques are now extensively utilized in airborne systems, data exchanges being performed in digital formats through busses, an adaptative, or at least reprogrammable, processing being more and more required. In this way the advent of microprocessors has been the starting point for an overwhelming digital processing.

- Nevertheless digital bandwidths cannot be increased indefinitely as cost and power consumption are in direct relation. It can be advantageous to utilize new CCD and SAW technologies to solve specific operations with less cost, size or power consumptions. A typical example is a SAW transversal filter designed for $T = 1 \mu S$ and $B = 500$ MHz ($TB = 500$). A digital equivalent would require 10^3 products to be summed at a sampling rate of 1 GHz = a computing rate of 10^{12} products per seconds.

- SAW and CCD are not competing and have their respective areas of applications as a consequence of their relative physical characteristics:

- SAW operate up to frequencies of 1 GHz, CCD being normally limited to frequencies of operation around 2 KHz and very well suited for low frequency operations.

— Storage time are of magnitude 100 μs for SAW and 1 second for CCD with respective bandwidths 600 MHz and 10 MHz, the product TB for both being in the area of $10^3 - 10^4$.

— There is a marked parallelism in the signal processing functions possible with CCD and SAW devices. This is due to the fact that both are analogue delay lines where the signal is accessible all along the lines. However SAW have a fixed propagation velocity, the passage through a CCD being dependent upon a time clock rate.

— The main success of SAW devices is based on variants of transversal filters = IF bandpass filters, dispersive delay lines for waveform generation and matched filtering, chirp filters for fast-acting spectrum analyser.

— CCD applications of a considerable interest are found in demultiplexing, time delay and integration, correlation. For the last functions must be noted a successful operation of a hybrid analogue/digital correlator in which analogue samples in a tapped delay lines CCD connected or not to summing lines, the connections being determined by the individual bits of a digital reference function. However, the real gain of such a combination over digital processing is not yet obvious and has to be proved by a gain on the power consumption.

— As far as delay lines are concerned, no firm requirement has yet emerged for CCD, except when advantage is taken of the freedom to change the clock rate as a means of adapting the time scale of a signal.

— Such an adjustment could lead to a large market of CCD adaptive tunable filters in citizen's band radio receivers.

— As far as imagery is concerned, CCD is ideally suited as a self-scanned solid-state imager, a very good example being a 400 x 400 area imager developed for the Jet Propulsion Laboratories.

— Monolithic and hybrid IRCCD have been built which are now performing very well.

Session 2 — Session 2 was devoted to technology and devices.

— Materials and fabrication processes for improving the performances of SAW devices have been presented in papers 2.1. – 2.2. – 2.4. – 2.6.

Paper 2.1. gives an interesting view of the materials selected to improve two classes of SAW devices:

a — Temperature-stable, broadband, low insertion-loss devices.

b — High-frequency devices.

Lithium niobate and aluminium nitride are, respectively, the substrate material best adapted.

Paper 2.2. presents a fabrication process using an electron-beam pattern generator to design SAW delay lines operating up to 1.3 GHz with a 3 db bandwidth of 500 MHz and a delay time around 10 microseconds.

Paper 2.4. is a description of a new type of pulse compression filter using the reflective dot array which has the advantage of being part of the same mask and metalization as the interdigital transducers. A linear filter operating at 60 MHz with a bandwidth of 20 MHz and a time delay of 10 microseconds is presented.

Paper 2.6. Describes a new technique for achieving frequency control of SAW oscillators with an external magnetic field – The accoustic propagation medium is a Gd — YIG film grown on a G.G.G. substrate. A frequency change of 1800 ppm has been obtained around 339.5 MHz for a magnetic field of 70 Oe. That opens the way to tunable SAW oscillators.

— The adoption of a 2-phase technology for fabricating charge coupled devices has resulted in a significant reduction of complexity in the driving requirements. This is the subject of *Paper 2.7.* which gives the performances of a two phase CCD fabricated with a coplanar, N — channel, double-level polysilicon gate process.

— Devices using SAW and CCD techniques are presented in papers 2.3, 2.5, 2.8.

Paper 2.3 is a description of a SAW pulse-compression device giving a TB = 500 with a bandwidth of 100 MHz and a dispersion of 5 microseconds.

Paper 2.5 gives some results on the influence of acceleration on SAW oscillators. An example is given at 124 MHz, the sensitivity being 48 HZ/g for an acceleration perpendicular to the propagation surface.

Paper 2.8 is devoted to the design of a CCD programmable transversal filter and correlator.

A parallel analogue convolver represents the best combination of speed and packing density at the expense of accuracy, as compared with a parallel digital convolver which represents the best combination for speed and accuracy at the expense of power and size.

— This is an excellent demonstration of the choice case-by-case of the technology best suited to solve a specific problem. When accuracy is not the main parameter but space and weight are at premium, analogue converters can find an interesting application.

— Complementarity between CCD and SAW techniques is demonstrated in a specific application of a high-speed signal digitizer presented in paper 2.3.

— The session was closed by *paper 2.10* which provided a comparative sensitivity study of CTD, SAW and digital processors. Processors considered include filters, spectrum analyzer and basic mathematical operations. This paper is intended to identify and model sources of error and parametric variations, and determine their effect on performance deviation from ideal performance.

Session 3 — Session 3 was devoted to functional modules/subsystems.

Papers 3.3, 3.6, 3.7, 3.10, present SAW technique as applied to oscillators, band pass filters.

Paper 3.3 describes a new type of SAW device, the SAW resonator using two possible materials, YZ Lithium Niobate and ST quartz, which are emerging with the ability of obtaining frequency stabilizing elements for oscillators and narrow bandpass filters.

Paper 3.6 describes the development of an acousto-electric diode memory correlator. Schottky diodes are charged in a few nanoseconds and have storage time of milliseconds allowing to keep in memory the phase and amplitude of surface wave signals. Other signals launched in the SAW delay live are correlated with the stored charge pattern.

Paper 3.7 is related to a memory correlator using an array of PN diodes coupled to a piezoelectric surface. This new device is a typical application of SAW technique perfectly adapted to the design of a linear filter used for correlation purposes.

— The bandpass is expected to reach 100 MHz for duration of signals around 20 microseconds, with a BT around 2000. The frequency of operation can be up to 300 MHz.

Paper 3.10 is an application of SAW filters to I.F. amplifiers used in a radar. Two different types of difficulties are overcome with SAW technique: one is to obtain a relative bandwidth of 0.3% which is very difficult to realize with conventinual LC filters, and the second is to obtain a rectangular bandpass with small ripples in the pass band, and sidelobes lower than 30 db.

— Papers 3.1, 3.2, 3.4, 3.8, 3.9 present CCD technique as applied to buffer memory, transversal filters, matched filters, arithmetic capabilities.

Paper 3.1 describes a buffer memory utilizing a CCD serial storage, which can interface with widely varying data rates between 150 KHz and 4 MHz. The memory is a 4 Kbits randomly accessible.

Paper 3.2 is devoted to a transversal filter where the weightings coefficient can be altered in real-time through a microprocessor used to control the tap weightings.

Paper 3.8 describes a new logic family, referred to as digital charge coupled logic. The family consists of AND, OR, exclusive OR, INVERT, HALF and FULL ADDERS, shift registers. Arithmetic capabilities are utilized in filtering applications: correlation, convolution, and fast transformers. Several applications of DCCL arithmetic are shown including the design of forward and reverse HADAMARD transforms.

Session 4 was devoted to imagery. Unfortunately, due to the high degree of classification of IR imagery, this session has been less promising than expected. Nevertheless it is felt that CCD have found their best applications in imagery technique.

— Two papers were presented on IR imagery. Paper 4.1 gives an interesting view on a new generation of FLIR using a bimensional array of detectors where the advantages of CCD or CID processing are clearly demonstrated by reducing the number of amplifiers normally required in the present state-of-the-art.

Paper 4.2 provides a general review of imaging sensors using CCD technology which are grouped in two major classes: Monolythic and hybrid IR CCD. Operations of both technologies have been analyzed.

Difference applications of CCD to imagery techniques are shown in papers 4.3, 4.6, 4.10, 4.11.

Paper 4.3 describes an optimum solid-state line scanner using a combination of photodiodes and analogue shift registers. This combination, referred to as charge coupled devices photodiodes (CCPD), possesses all the advantages of the photodiode detectors and the analogue shift register read-out.

Paper 4.6 is devoted to an interesting application of an array of CCD sensors to identify linear obstacles such as wires which can be extremely dangerous for nap of the earth operations of helicopters.

— The CCD photosensors are used in a gated TV technique with an illumination provided by Ga As diodes.

Paper 4.10 concerns a CID solidstate video sensor which produces a HADAMARD transform of an incident optical image.

— This paper proves the feasibility of real-time coding of an optical image directly on a CID imager. This development is the first step to the realization of a miniature, low-power, jam-resistant video data link suitable for mini RPV, guided weapons and cannon-launched guided projectiles applications.

Paper 4.11 describes a high-performance CCD linear imaging array.

Session 5 — The final session was devoted to signal processing. Excellent papers were presented showing applications of CCD and SAW techniques in signal processing.

— The first three papers 5.1, 5.2, 5.3 describe applications of SAW, CCD, and a combination of CCD and SAW to the chirp — Z transform.

— SAW chirp — Z transform as applied to a pulse doppler microwave radar had a 40 db dynamic range and frequency determination achieved with an accuracy of 0.01% for a total process time of 50 microseconds.

— A combination of CCD, as a sampling memory and a SAW dispersive delay line has permitted the design of a low-cost, low-power, small-size Fourier Transform signal processor.

— The CCD chirp — Z transform has shown limitations relative to the digital fast Fourier transform. Nevertheless, applications can be found in video bandwidth reduction, speech processing, FLIR, sonobuoys, when low cost is a dominant design specification associated with modest performances and sufficiently high volume.

Typical applications where SAW technique is perfectly adapted are described in papers 5.4 and 5.5.

Paper 5.4. describes an acoustic convolver designed for several-signal radar processing and more particularly to synthetic aperture radars. An interesting comparison with digital convolvers shows a number of advantages for SAW technique, one of the most important being to suppress the need for a rough phase quantification needed in digital techniques which is a cause for quality degradation.

Paper 5.5. is devoted to pulse compression application in a high-performance radar. A TB of 5120 has been realized with a bandwidth of 512 MHz and a pulse width of 10 microseconds. The frequency of operation is 1 GHz. Range separations between targets has been as small as 4 nanoseconds.

— Although the advantage of CCD technique is not so obvious, *Paper 5.7.* presents an application of CCD delay line to MTI clutter rejection. The results are very promising as far as cost, size, power dissipation are concerned. Expecting an increase in the dynamic range available, it seems that CCD technique could be a good candidate for this type of application.

CONTENTS

SESSION V — SIGNAL PROCESSING

STATE-OF-THE-ART OF CCD AND SAW TECHNOLOGIES

Lewis T. Claiborne
Texas Instruments Incorporated
Dallas, Texas 75222

SUMMARY

This paper presents a survey of the status of component technologies for charge coupled devices (CCD) and surface acoustic wave devices (SAW) in terms of device performance and application to subsystem signal processing functions. The performance will be discussed for imaging and image processing as well as analog and digital signal processing. Recent developments in SAW technology will be mentioned including resonator filters and transform processing.

I. INTRODUCTION

A generalized statement of systems functions addressed by these component technologies will be presented in order to provide a framework for subsequent discussions of key topics of the symposium, (i.e., signal processing, including memory applications, and imaging, both visible and IR). Specific components which serve these general systems' needs will be reviewed. Examples will be presented of well developed components which are representative of the present state-of-the-art as opposed to one-of-a-kind research results. Time will not permit a complete exposition of all the possible devices which have been described in the literature. Many important devices which have been omitted here will be covered in other talks at this symposium. The conclusion of this paper is a set of tables presenting current and projected performance for each type of component.

II. GENERALIZED SYSTEM REQUIREMENTS

CCD and SAW components have many significant roles to play in today's systems. In order to understand the implications of the device performance in the context of fundamental systems requirements, consider a specified block diagram of a generalized system which could apply to radar, communications, or imaging systems as shown in Figure 1. Such a system has an input for signal or data acquisition followed by preprocessing, some type of data processing, and an interface which leads to a display, data link, or transmitter.

To illustrate what is meant for a generalized system concept consider the simple example of a TV system. First, a TV receiver has a front end consisting of a tuner with prefiltering, followed by preprocessing in the form of the complete tuner function with the IF. The data processing section includes color, video, and sound demodulation. The interface in this case could be simply a buss going to a display which is usually a CRT. Alternatively, of course, a TV transmitter system can have a CCD camera for the input with the processing roles reversed, and the output would be a TV transmitter. To apply this structure to systems of general interest to this audience consider first a general radar system. In a radar system there are a number of places which the CCD and SAW components can play major roles, for example, in front end filtering and preprocessing through pulse compression. Frequently there is a requirement for doppler processing followed by multiplexing (or sometimes memory) and then radar display. A spread spectrum receiver system also requires front end filtering. In general, there is a need for system syncnronization and, perhaps, a need for a fast hop frequency synthesizer. In the preprocessing area there is a general requirement for prewhitening and matched filtering. Data processing can be digital data processing or perhaps transform processing. The interface function would frequently require memory. Imaging systems in a similar context could use charge coupled device imagers, intensified imagers, infrared CCD's and IR detectors. Frequently, bandwidth compression is a desired preprocessing step. Also, frequently, image correlation and target acquisition are functions for the processing section followed by memory, scan conversion, and multiplexing leading to display system. These very generalized system comparisons are intended to illustrate certain functions which are required for either components or subsystems and for which the CCD and SAW are well suited. Most of these systems, for example, require various types of bandpass filtering. There is a need for matched filtering, correlation, pulse compression, and the prewhitening for spread spectrum communications, as well as more sophisticated subsystems such as the doppler processor. Analog and digital memory are frequent prime requirements. The ability to perform scan conversion to reformat data consistant with display requirements is another very important application area.

In order to show how these types of general requirements lead to some specific applications of these technologies, in Figure 2 the same breakout of system functions is used to illustrate where some of these requirements occur. To make this comparison meaningful in terms of the component technologies that are to be considered at the symposium, the frequency ranges of interest are plotted across the bottom of Figure 2 along with words to illustrate where imagers, SAW filters, and CCD devices provide the necessary functions. This is a very qualitative indication of the position of each of these component technologies and is simply intended to illustrate the wide range of functions and the large number of sockets that these components can fill.

III. SIGNAL PROCESSING

The discussion of signal processing will be broken up into some general classes of functions. First will be discussed the transversal filters, recursive filters, signal transform devices, adaptive filters, various types of correlators, and finally convolvers and combined SAW/CCD signal processing functions. Figure 3 shows the surface acoustic wave transversal filter. For historical reasons this device is being described first. The surface wave filter consists of two or more interdigital transducers which are metal electrode-arrays deposited on a piezoelectric substrate. The input transducer converts electrical signal into elastic signal in the substrate in the form of a Rayleigh wave which propagates along the surface. Interdigital transducer taps or output transducer can be placed at any point along the propagation path, and the amplitude and phase of the output taps can be arbitrarily specified by the geometry of the electrode pattern. This device is a particularly simple example of a transversal filter structure. The impulse response of the filter is specified in terms of the transducer geometry and the

output from the filter is a convolution of the input signal with the impulse response of the device. The SAW filter is inherently an analog device with large dynamic range possible. It is also possible to obtain good efficiency for the transduction from electrical to acoustical signal.

The CCD is similarly a very simple semiconductor device for which surface electrode potentials are used to store and manipulate charge along the semiconductor surface. Analog signal in the form of charge is stored in potential well underneath these electrodes. The structure shown in Figure 4 is a three-phased CCD with a coplanar electeode structure. This sort of device is ideally suited for clocked delay lines which can be used for analog filters. These devices can also be used for analog and digital memory and for self-scanned solid state imagery depending on the method used to introduce the charge into the device structure. The CCD relies on the clocking of the potentials of the electrodes to create a delay line. The manner in which such a delay line can be used to create a sampled data filter is shown in Figure 5. Discrete delay stages and the means for tapping the delay along the device can be provided. The mathematical equations which represent this operation are also shown. One method of sampling the conventional CCD employs split clock electrodes and differential sampling of current of one of the clock phases as shown in Figure 6. The current in each line depends upon the charge under each electrode. As with the surface wave device, geometries can be used to control the impulse response of the CCD filter. It should be noted here that the CCD filter lends itself ideally to integrated circuits incorporating transversal filters with other signal processing circuitry.

A comparison of the CCD ans SAW technologies reveals certain differences which lead to choices based on advantages for various system applications. For example, the mode of operation of surface wave device is passive while that of the CCD is active. The SAW is in general a fixed delay line while the CCD is variable with the clock rate controlling the delay time or the rate of data propagation through the device. The SAW has a continuous sampling while the CCD is a sampled data filter. SAW's have typically good temperature stability based on the inherent properties of the piezoelectric substrate while the CCD stability can be controlled by a master clock which controls the clocking of the delay line. The surface wave device has a large dynamic range (typically > 100 dB) while that of the CCD is also proven to be quite large (> 70 dB) and is quite useable for many applications. Taking the comparison one step further and ignoring the details of the specifics of the two technologies, many similar functions can be realized. Figure 7 gives a simplified view of the operating bandwidth and processing time of the two types of devices. The surface wave device can typically provide bandwidths of up to 1 GHz and operates generally between 10 MHz and 1 GHz. The CCD is basically a MOS device which in transversal form required peripheral circuitry which can in fact limit the upper frequency of operation to $\sim$ 2 MHz. However, the clocking can be also performed slowly and the CCD's can operate down to very low frequencies. These two devices are quite complementary in the range of operation and in comparisons where the same function is to be provided either in VHF or UHF as compared with RF frequencies and lower frequencies there is an obvious choice between the surface wave and the CCD device. Time bandwidth products for the two devices range in the neighborhood of 10^3-10^4, perhaps even higher in the case of SAW devices. As will be mentioned subsequently, the combination of surface wave and CCD technologies can provide a quite large time bandwidth products subsystem.

A.   Transversal Filters

Some examples will now be shown of some of the applications of components which are readily available today. One of the first successful applications of SAW technology was in the area of pulse compression filters for radar systems. Shown in Figure 8 is a typical result which is several years old. Time bandwidth products of greater than 100 and sidelobes suppression of greater than 30 dB is readily available at IF frequencies which lie in the VHF and UHF bands. More recent work using reflective array correlators (RAC) has demonstrated time bandwidth products up to 10^4 and superior phase control. Matched filters for pulse compression and code/decode functions have been demonstrated for CCD as well.

Examples of transverse bandpass filters are shown in the next three figures. First, good out-of-band rejection is illustrated by a SAW filter sintered at 287 MHz. Rejection of greater than 70-80 dB is shown across the spectrum from dc to 1 GHz. SAW filters of this type have traditionally required moderate insertion loss on the order of 20 dB to suppress triple transit echos and distortion effects. This fact has limited their application to IF functions. It is now possible to achieve low insertion loss with low distortion with moderate out-of-band rejection using unidirectional techniques as shown in Figure 10. As shown here, filters for front end applications are also possible, and insertion losses of less than 1 dB have been demonstrated. Transducer structures of this type require multiphased arrays and bi-level metallization systems for fabrication, and techniques for manufacturing the devices have been developed. This technology should make low loss filters a standard component of many system applications. CCD bandpass filters have also been demonstrated and are now being exploited for real systems applications. For example, Figure 11 shows the bandpass of a filter IC design for a citizen's band radio application. These CCD filters are well suited for RF/IF filter requirements and can be expeceted both in commercial and military applications.

B.   Recursive Filters

In addition to the transversal filters described above, recursive filters are now possible for both SAW and CCD. For the SAW technology a particularly important device is the trapped energy mode resonator. A schematic illustration of the SAW is shown in Figure 12. A cavity is formed by distributioned reflectors deposited or cut into the surface of the piezoelectric crystal substrate. An interdigital electrode transducer can be used to couple to the cavity which is formed. This device is analogous to the bulk wave resonators, but has the advantage that it can be fabricated for frequencies up to L-band. While the conventional bulk wave resonator structures are normally limited to frequencies of $\sim$ 10 MHz, the surface wave resonator can also provide Q's of greater than 10^4, and both frequency control and narrow band filter functions can be provided in the VHF and UHF bands. Figure 13 illustrated the performance of a typical SAW quartz resonator. Using devices of this type, multipole filter structures can be realized. The stability of the surface wave resonator with temperature is that of substrate. For example, the so called ST cut of quartz for surface waves is within a factor of $\sim$ 3 of the stability that would be obtained with an AT cut crystal for bulk wave resonators. The short term stability of SAW resonator devices is quite good. For example, on the slide the Allen variance is plotted versus the averaging time for a surface wave resonator oscillator. It can be seen that the variances approaching 10^{-10} can be obtained for averaging times in the neighborhood of 10 msec. For the CCD recursive filter the modes of feed forward and feedback can both be

used to provide zeros and poles. Recursive filter theory can be applied for various types of filter designs.

C. Transform Subsystems

Another important application area for these technologies has been in the realization of signal transform functions. Recognizing that (1) arbitrary impulse responses can be obtained and that (2) these can be coupled with convolution operations has lead to the implementation of subsystems for various algorithms. A powerful technique which has been demonstrated is the chirp Z-transform which can implemented in sample data form as shown in Figure 15. Here, a time ordered sequence of Fourier coefficients can be obtained for finite time sequency input waveform. Thus, it is possible to build subsystems which can transform signals from the time domain into frequency domain and through the inverse transform to go back from frequency to time domain in real time. The building blocks for this illustration are simply chirp transversal filters which are used for signal premultiplication and postmultiplication operations with a third filter providing convolution with a chirp filter. To see how this actually works in practice, consider the analysis of Fourier coefficients of a square wave as shown in Figure 16. An input 200 Hz waveform is shown in the upper trace A and the output spectrum going up to the 19th harmonic is shown in the lower traces of B,C,D. It is clear that the time ordered sequence of Fourier coefficients is just as expected with the odd harmonics falling off with the order. This technique has been applied to several system requirements. An airport surveillance processor is just one shown here in Figure 17. The processor is used to separate radar returns for both doppler and position by sorting the signal into range bins and by analyzing the signal in terms of the doppler shirt by taking Fourier transform.

Similarly, real time signal transforms are possible with surface wave devices. For example, a chirp transform again based on individual transversal filter components is shown here to implement a 120 point transform for a 2 μsec interval. The surface wave device is operating on a signal carrier and thus we see pulses of cw where the time order depends on the frequency of the input signal. There are a number of applications for this technique for the VHF and UHF range such as spread spectrum receivers, with the pre-whitening, spectral analysis, and so on. Further, these subsystems can be used to perform higher level functions. For example, in order to perform matched filtering in a manner similar to that in which digital process is performed, the correlation of two waveforms can be achieved by taking the inverse transform to obtain the matched filter output. Examples of this technique and operation will be shown for a simple PN sequence. Thus, the ability to take transforms of waveforms provides alternative tools for doing adaptive data processing. The operations that are possible include various types of bandpass filtering with amplitude weighting in the frequency domain, various types of matched filtering, and so on. The list of applications is actually quite large and this technique that is just beginning to be realized for future systems application.

D. Adaptive Correlators and Convolvers

There are other alternative techniques for providing adaptive filtering for both CCD and SAW technologies. Shown in Figure 20 is an illustration of the correlation of two analog waveforms using the sampled data technique. Here an input waveform is put through succeeding delay stages and sampled at each point. These samples can be multiplied together with similar samples from the reference function and the outputs can be summed in parallel to provide an analog/analog correlator. This format is particularly well suited for the charge coupled device. There are many other variations, but time does not permit a survey at this point.

Similar functions can be performed for SAW devices through the convolver structure. This is an interesting hybrid device in which the piezoelectric substrate is coupled to an adjacent semiconductor. Signals which are introduced at either end of the substrate can be sampled and the parallel products can be summed along the propagation path through the interaction of the surface wave fields with the semiconductor. This type of device has recently been shown to be practical on a laboratory scale and is shown in Figure 21. The assembly of the surface wave piezoelectric substrate and the semiconductor, in this case silicon, can be accomplished with fairly straightforward packaging techniques, although this type of device is certainly not as well developed as the transversal filter which has been discussed earlier.

E. SAW/CCD Processors

A final type of processor which can be provided by these technologies is actually a combination of the SAW/CCD device technologies. In many instances in data processing it is desirable to provide a two dimensional Fourier transform. The surface wave device because of its high bandwidth capabilities can provide very rapid real time one-dimensional Fourier transform. The CCD because it is a device which allows one to manipulate analog information in an array form is ideally suited to provide a corner turning memory (or scan converter as it is sometimes called). The operations which can be achieved through the use of one-dimensional transforms plus CCD array manipulation immediately suggests that two-dimensional Fourier transforms can be performed using the techniques which are conventionally used with digital processors. High speed surface wave input Fourier transforms can be performed on data to provide a one-dimensional trans-form. The data can then be read into an array and this array can then be addressed in a perpendicular direction to provide row in-column out in such a way that the second dimension Fourier transform can be provided with a second high speed surface wave Fourier transform. The combining of these technologies makes it possible to provide very fast two-dimensional Fourier transforms. For example, it is possible to take an array which is roughly 1024 x 1024 and provide a real time 2-D Fourier transform in the time of perhaps 5-10 milliseconds.

IV. MEMORY DEVICES

It is beyond the scope of this overview to go into the digital memory applications of the CCD technology. However, for completeness, one slide will be shown here which discusses what is currently available and what is projected in terms of semiconductor memory through the CCD technology. For example, a 65,536 bit CCD memory part is currently available by several manufacturers. Typical performance by such a device (TMS 3064) provides a maximum data rate of 5 megabits per second, minimum data rate of 1 megabit per second, average latency time of 400 microseconds, power dissipation would be typically 300 mw max. It is projected that by the year 1979-80 it will be possible to obtain a factor of 4 higher bit capacity chips (262,154 bit chips). The surface wave technology can also lend itself to various types of memory technology.

The applications directly to digital memory are not as clear at this point in time. However, the ability to provide long delay times has been demonstrated. Up to millisecond delay times can be achieved with the use of amplification in the surface wave propagation path. There may be roles for analog memory for the surface wave technique as well, particularly where high data rates are desirable. Again this technique is beyond the scope of this particular survey.

V.   IMAGING DEVICES

As has been aluded to earlier, CCD is ideally suited as a self-scanned solid state imager. Charge can be introduced into the various locations in the CCD array from incident light by integrating for a finite period of time. An area array is particularly well suited for the imaging applications. Some imager formats are shown in Figure 22 for typical imager configurations. The full frame area array is shown first. In this case, charge packets representative of local light load are stored in a matrix during an integration time. The signal is read out by shifting a line at a time into the output shift register. This type of operation requires a shuttering of the camera or a large integrate/readout time ratio. In the frame transfer mode, one section of the imager collects the image which is transferred rapidly into a memory section. Readout of the memory is accomplished while the new image is being collected. Another technique which has application both in visable and infrared for a moving image is the time delay and integrate mode of operation. The image is integrated from pixel to pixel in a synchronous manner. This time delay and integrate (TDI) technique provides a continuous input and continuous output. Other techniques for imaging include the GE CID and the Fairchild interline transfer scheme. The system tradeoffs in terms of performance, area, frame rate, system condiguration, etc. are applications dependent and determine which format is chosen.

To illlustrate a real CCD imager, Figure 23 shows a 400 x 400 area imager which was developed for the Jet Propulsion Laboratories. This is an example of a full frame imager. The particular technology used here is backside illumination of a thin CCD area array. (The light enters from one side and the electrodes are on the other). This is a large bar being approximately 1 centimeter x 1 centimeter. There are several imagers of similar size which are available commercially today. These represent perhaps the present state-of-the-art of approximately 160K to 185K elements. Examples of imager performance for two different size area arrays are shown in Figure 24 where a 1 MHz readout rate is used for a 400 x 400 area array and a 160 x 160 area array. These two designs have been useful for imaging development. The 400 x 400 imaging array has high resolution obtained because of the large number of pixels. The detail can be seen in the hair and eyes. It is possible also to obtain low defect imagers even with such large arrays as can be seen from this photograph. A smaller number of resolution elements in the second array still provide a fair amount of information.

A limiting feature of the CCD imager is the so called dark current which tends to block out the information contained in the potential wells as the wells fill up with charge. However, rather long integration times can be achieved as shown in Figure 25 where at room temperature storage times on the order of many seconds can be achieved with still relatively good imager performance. The storage time is limited by variations of the dark current pattern which accumulates for the full period as indicated. Cooling these sensors, of course, would imply longer storage times and also would imply that one could obtain better sensitivity since longer integration times could be used.

The CCD imager has some desirable features for various types of image process systems. For example, a photograph of Uranus was obtained using a ground based telescope and the 400 x 400 element area array shown previously.* The work here provides detail not formally observed using other sensor techniques. The photograph shows, in particular, an example of the application of long integration time. The photograph was taken using an imager cooled to -40°C and is the result of 7 frames that have been processed by a computer. Each of the frames is obtained using two 2 minute exposures. Actually, this photograph provided detail which was not previously available. First, the polar region is observed using filtering for methane absorption at 8900A which shows the polar region as a dark area. This is the first feature that has ever been observed from a ground based telescope fo Uranus. The second key point is that it was possible to get a good measure of the diameter of Uranus which led to new estimates for the density of this planet. The desirable features of the CCD here were the linear input/output characteristics and the fact that the individual pixel locations is very accurately known.

Another technique for obtaining good low level light performance is shown in Figure 26, i.e. the use of the electron bombarded CCD imaging tube in which the insident light strikes a photocathode and accelerated photoelectrons are used to obtain gain. The backside technology is again used here with the electrons impinging on the back of the CCD array. The use of gain before the sensor relaxes requirements for dark current noise level and cooling. The gain is obtained by accelerating the photoelectrons to several thousand electron volts and re-imaging on the silicon backside surface of the imager. The net result is that the intensified CCD is photon noise limited to signal level approaching 1 photon. Figure 27 shows the signal to noise ratio obtained versus the face plate illumination for direct view of silicon and intensified CCD using the GaAs photocathode. The curves show the improvement in signal to noise ratio that is obtained using photocathodes. It should be pointed out that the direct view silicon here is assumed to be at room temperature. By cooling the silicon imager, it is possible to move the curves down so that at low light level illuminations better signal to noise ratios are obtained.

VI.   CONCLUSION

The present and projected characteristics of the CCD and SAW devices are summarized in the following three tables. From the brief outline of systems requirements and device functions presented in this paper, it can be seen that CCD and SAW components are subsystems have a key role to play in future radar, communication, and imaging systems.

*B. Smith, Department of Planetary Sciences, University of Arizona, Tucson

SAW DEVICES

Characteristics	Present	Projected
General		
Center Frequency	10-500 MHz	10-2500 MHz
Temperature Stability	∿1 ppm/°C	∿0.1 ppm/°C
Min. Insertion Loss	2 dB	0.2 dB
Transversal Filter		
Max. Fractional Bandwidth	50%	100%
Max. Frequency Sidelobe Suppression	50 dB	70 dB
Max. Time Length	120 μsec	250 μsec
Amplitude Ripple	0.1 dB	0.02 dB
Phase Ripple	±2°	±0.5°
Resonator Filter		
Max. Unloaded Q 100 MHz*	40,000	80,000
Spurious Suppression	-25 dB	-40 dB
Aging	10 ppm/yr.	0.5 ppm/yr.
Complexity Level in Monolithic Form	2 port, 1 pole	2 port, 4 pole

*Varies inversely with frequency

TABLE 1

CCD TRANSVERSAL FILTERS

Characteristics	Present	Projected
Bandwidth (W)	2 MHz	50 MHz
Time Delay (Td)	100 Msec	1 Sec.
TdW Product	400	10,000
Dynamic Range	60 dB	100 dB
Stop Band Rejection	-40 dB	-60 dB
Harmonic Distortion	-50 dB	-70 dB

TABLE 2

CCD IMAGERS

Characteristics	Present	Projected
Array Size	400×400	800×800
Resolution Element Size	25 × 25	10 × 10
Dark Current Levels (300K)	3 nA/cm^2	0.5 nA/cm^2
Noise Equivalent Signal	$1.4 \times 10^{-4} \frac{1um}{M^2}$ (Clear Night Sky)	$2.3 \times 10^{-5} \frac{1um}{M^2}$ (Overcast Night Sky)
CTE	0.9999	0.99999

TABLE 3

VII.   BIBLIOGRAPHY

The CCD and SAW technologies are represented by a large body of literature. Rather than providing a detailed, but incomplete list of references, a guide to some recent survey material is presented below.

Books

Collins, J.H. and Masotti, L., 1976, Computer-Aided Design of Surface Acoustic Wave Devices, Elsevier Scientific Publishing Company, Amsterdam

Melen, R. and Buss, D., 1976, Charge-Coupled Devices: Technology and Applications, IEEE Press, John Wiley and Sons, Inc., N.Y.

Conference Proceedings

Proceedings of the International Specialist Seminar on "The Impact of New Technologies in Signal Processing", September 1976, Aviemore, Scotland, IEEE Conference Publication No. 144

Proceedings of the Third International Conference on "The Technology and Applications of Charge-Coupled Devices", September 1976, University of Edinburgh

Journals

Barbe, D.F., "Imaging Devices Using the Charge-Coupled Concept", January 1975, Proc. of the IEEE, Vol 63, No. 1, pp. 38-67

Holland, M.G. and Claiborne, L.T., "Practical Surface Acoustic Wave Devices", May 1974, Proc. of the IEEE, pp. 582-611

CCD Technology and Review of CCD '75, Microelectronics Journal, December 1975, Vol. 7, No. 2

Special Issue on Charge Transfer Devices, IEEE Journal of Solid State Circuits, February 1976

Special Issue on Surface Acoustic Wave Devices and Applications, Proceedings of the IEEE, May 1976

GENERALIZED SYSTEM REQUIREMENTS

	Signal/Data Acquisition	Preprocessing	Processing	Interface	Display or Data Link
TV	Tuner/Prefilter CCD Camera	Tuner/IF Adaptive Equalization	Color Video and Sound DEMOD	Bus	Display Transmission
Radar	Front End Filter	Pulse Compression	Doppler Processor	Multiplex Memory	Display
Spread Spectrum Receiver	Front End Filter System Synchronization Fast Hop Synthesizer	Prewhitening Matched Filter	Digital Data Processor Transform Processor	Memory	
Imaging System	CCD ICCD IRCCD IR Detectors	Bandwidth Compression	Image Correlation Target Acquisition	Memory Scan Conversion Multiplex	Display

1

GENERALIZED SYSTEM FUNCTIONS
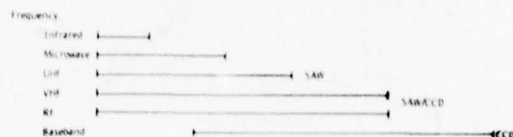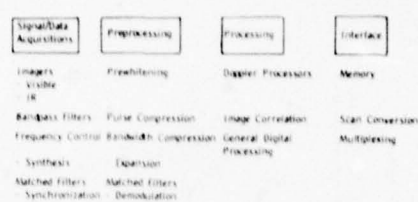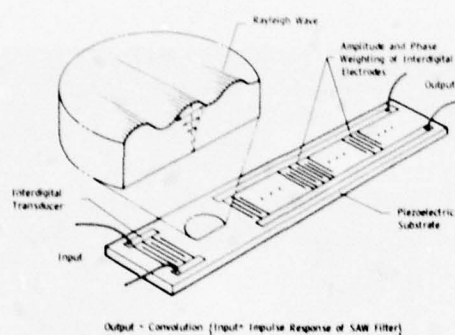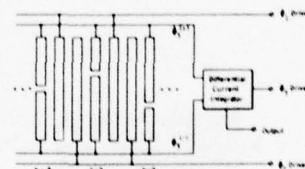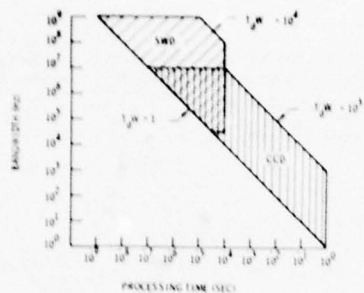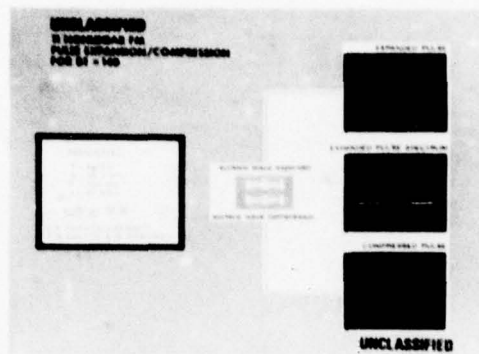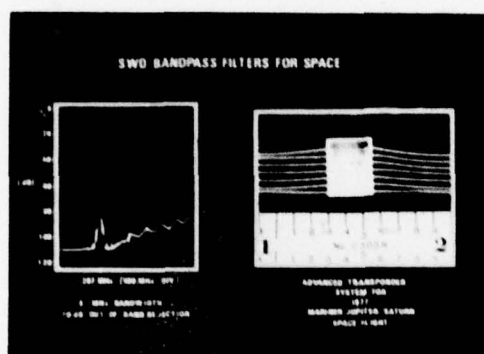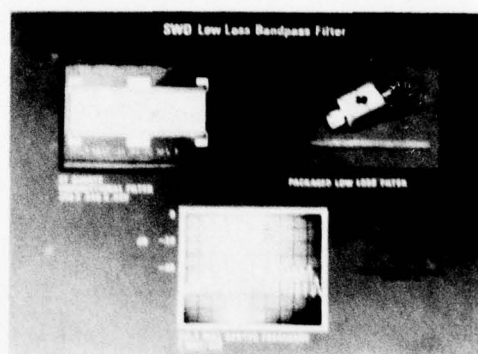
Signal/Data Acquisitions	Preprocessing	Processing	Interface
Imagers Visible IR	Prewhitening	Doppler Processors	Memory
Bandpass Filters	Pulse Compression	Image Correlation	Scan Conversion
Frequency Control	Bandwidth Compression	General Digital Processing	Multiplexing
Synthesis	Expansion		
Matched Filters Synchronization	Matched Filters Demodulation		

Frequency

Infrared
Microwave
Uhf — SAW
Vhf — SAW/CCD
Rf
Baseband — CCD

2

SURFACE ACOUSTIC WAVE
TRANSVERSAL FILTER

Output = Convolution [Input* Impulse Response of SAW Filter]

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18
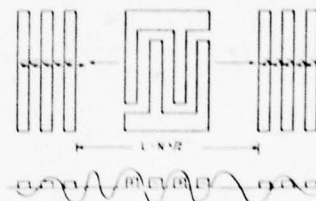
CHIRP TRANSFORM PROGRAMMABLE MATCHED FILTER

19
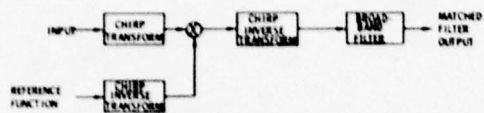
SITD/CCD SIGNAL PROCESSORS
CCD Autocorrelator/Correlator
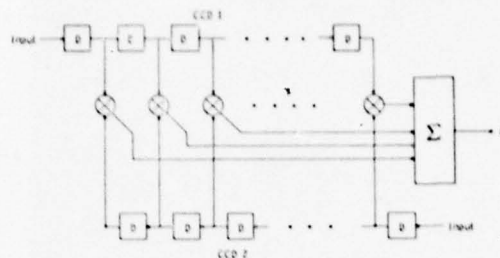
20

21

CHARGE COUPLED DEVICES TECHNOLOGY PROGRAM
IMAGER FORMATS
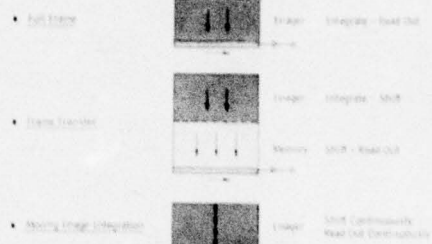
22
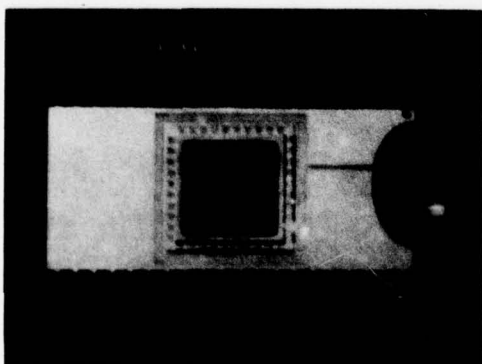
23

24
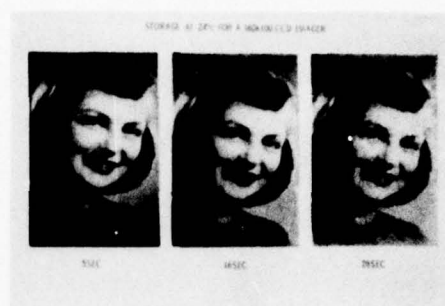
25
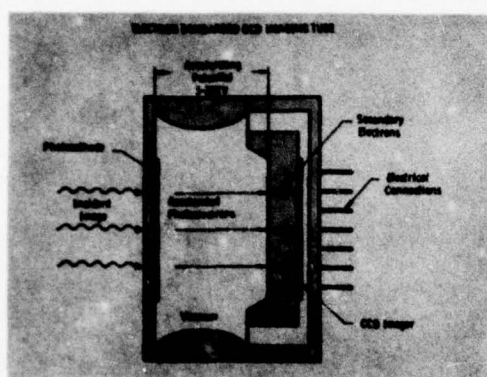
26
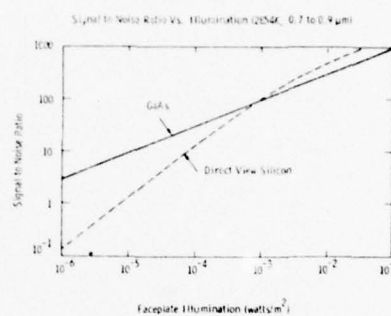
27

THE ROLES FOR CCD AND SAW IN SIGNAL PROCESSING

J B G Roberts
Royal Signals and Radar Establishment
Malvern, UK

SUMMARY

The recent emergence of two radically new methods for the real time processing of signals raised the important question of how they may best be used and how they relate to one another and to other techniques. We discuss the key factors determining the effectiveness of CCD and SAW in a range of application fields, assessing in which roles they are likely to be most appropriate and how they compare with the rapidly expanding range of digital devices. As well as contrasting the techniques in mutually exclusive situations, we show how it is often possible to complement one with another to good effect.

1. INTRODUCTION

In this paper we attempt to review the attractions of SAW and CCD techniques to the designer of signal processing subsystems, particularly in the present context of rapid advances in LSI digital technologies. A full account of such a wide field is not possible here and the more specialist reader is referred to the comprehensive discussions of the issues involved which is available in the proceedings of the recent Aviemore specialist seminar, (IEE 1976).
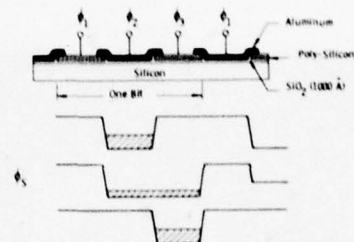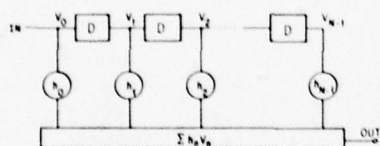
The approach here is intended to emphasize the achievements of SAW and CCD rather than to dwell on their possibilities for the future, which is the proper area of concern to other papers at this meeting.

2 CONSTRAINTS ON MODERN SIGNAL PROCESSING TECHNOLOGIES

At present the rate of advancement in the available performance (if not the rate of procurement) of electronic systems is probably higher than ever before, and is due increasingly to improvements in information processing techniques rather than developments in the more basic components of systems. The overwhelming trend over the last few years has been the enlargement of options available to system designers through the wide range of digital components whose speeds and costs have consistently shown exponential improvements by factors of almost two per year.

Against this dominating backdrop, alternative analogue signal processing techniques can emerge from the research laboratory into the real world only with extreme difficulty. This is firstly because of the immense investment backing provided by the growth of the computer industry and also because many systems (especially large ones) are now often conceived in a digital format whereby sampled data is assumed, interleaving of different data streams is required and adaptive or at least reprogrammable processing is advantageous.

It is now fair to claim that surface acoustic wave (SAW) devices have overcome this hurdle and established themselves as robust, compact devices with some capabilities unlikely to be rivalled by arithmetic processing 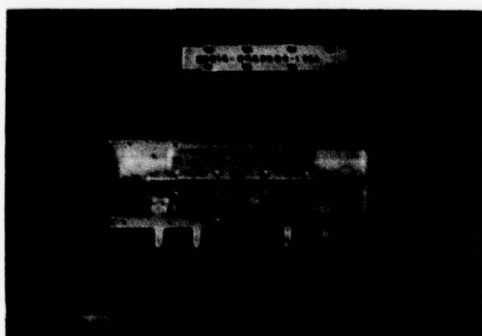for many years to come. However the position of the charge coupled device (CCD) is not yet so safe. The device technology is not so mature and its important signal processing parameters we shall see are rather closely in line with those of digital processors so that the balance of advantage is often determined by the secondary, but often vital, issues of power consumption, size and cost; the last of these being strongly sensitive to the volume of production.

Other technologies, for example magnetic spin waves, magnetic bubbles and coherent optical processors, while of interest in principle, have not yet received such widespread attention, probably because of their reliance on specialised materials or close mechanical tolerances. It is apparent therefore that a realistic appraisal of the likely impact for each type of device must pay attention to the commercial pressures brought about by the heavy commitment in industry to the technology of integrated circuits on silicon and by the needs of the consumer market for television, CB radio and video game components etc. A new technology requiring non-standard manufacturing methods or exotic materials and not able to earn its keep by satisfying a large scale commercial need is unlikely to supplant existing signal processing devices even in high performance military systems required in small numbers.

Both SAW and CCD are able to capitalize on existing industrial capabilities. SAW devices are fabricated almost exclusively on readily obtainable Lithium Niobate or quartz and use for the most part, standard single layer metallizations defined by conventional photolithography. Also the common need for carefully specified and reproducible IF filters in TV receivers is now providing a strong commercial stimulus for maintaining SAW technology. CCD uses the standard methods of silicon technology and its more specialised requirements for multi-level polysilicon gates and low dark-current material can be justified by the large scale markets, outside signal processing, for CCD memories and imagers. With both CCD and SAW therefore competitively poised from the viewpoint of device fabrication, we may consider their more fundamental attributes from the point of view of a system designer.

3 COMPARISON OF SAW AND CCD CHARACTERISTICS WITH THOSE OF DIGITAL COMPONENTS

3.1 Digital Processing

The very evident advantages of digital components include the convenience of a 'universal library' of commonly required simple functions, comprising for instance analogue-digital converters, multiplexers, stores, multipliers and adders. A dual second-order filter even is now available on a single chip. These permit the building up of more complex processors such as Fourier transformers and adaptive filters of

predictable and stable performance with no special development costs. The epitome of the all-purpose
programmable device is the microprocessor chip. Perhaps the greatest impact of the digital revolution
has been in the standardization of information handling methods. Data bussing, bit-serial or parallel
formatting, multiplexing and handshaking have all become commonplace operations and simplify the inter-
facing of a wide range of devices.

The universality and arbitrary accuracy of digital components is pre-eminent in decision making, control
functions and precise long term storage but from a signal processing point of view, these advantages may
be bought at some expense in power consumption, cost and limited bandwidth. For instance to perform
digital multiplication of samples at a rate of a few MHz typically requires several watts whereas a CCD
or SAW transversal filter may perform 10^9 or 10^{12} equivalent multiplications per sec with far less con-
sumption. Of course digital bandwidths can usually be increased by using proportionately larger volumes
of hardware but the cost and power consumption are also proportional. The speed, cost and size of digi-
tal hardware is, as we have noted, subject to exponential improvement with time but the same progress does
not seem to apply to the power/speed ratio. Table 1 illustrates the attainments of the major digital
technologies.

TECHNOLOGY	CMOS	TTL LS	ECL	I^2L
Density Gates/chip	100-1000	100-500	30-300	300-500
Prop. delay ns/gate	10-50	2-10	0.5-2	5-100
Power mW/gate	10^{-3} @ 1 kHz 1 @ 1 MHz	1-5	10-100	10^{-4}-10^{-1}
Note	Optimum Compromise	Low Price	High Speed	New Technol

TABLE I - Comparison of the major LSI technologies

Against this dominant theme of cheap, non-specialised, time-flexible digital building blocks with complex-
ity, power consumption and cost roughly proportional to the product T.B.D (signal storage time, signal
bandwidth, dynamic range) required in any application; the SAW and CCD techniques justify themselves in a
less general way by offering large bandwidths, cheap systems, miniaturisation or low power consumption in
specific cases where these are vital.

3.2   SAW Processing

The surface acoustic wave device is of interest in signal processing chiefly because of its following
characteristics:

   (i)   Its planar technology is realisable by standard IC methods using computer
          aided design.

   (ii)   Simple construction, cheap in volume production.

   (iii) Stable parameters, (time, frequency, phase, amplitude).

   (iv)   Very high bandwidth (hundreds of MHz).

   (v)   Accessibility of the signal at intermediate points within the device.

The device is basically an IF analogue delay line with a propagation delay of order 3 µs/cm with an opera-
ting frequency typically in the range 10 to 500 MHz and a fractional bandwidth up to 50%. However through
a detailed understanding of the device behaviour (diffraction, reflections, mass loading etc) it is possible
to design input and output transducers, couplers and reflecting structures to define very precisely a
variety of required impulse responses. Because the signal exists within the filter as a wave on the sur-
face of the substrate, multiple output transducers are possible, for instance to realise a tapped or variable
length delay line.

The established capabilities of SAW devices in signal processing systems largely depend on fixed-function
filters. This is not as limiting as might be thought, since several possibilities exist for controlling
system behaviour by exploiting the frequency dependence of filters. The determining characteristics of
SAW devices arise through the more-or-less fixed propagation velocity, the availability and convenience
of handling crystal substrates of no more than say 20 cm in length, and the sub-micron accuracy possible
in defining transducer structures by photolithography techniques. These result in typical storage or
delay times T below 70 µs and typical bandwidths in the range 15 kHz (~1/T) to several hundred MHz. The
parameters B and T cannot be chosen independently: for instance long delays are not available at the high-
est bandwidths and operating frequencies because of attenuation, and the precision of lithography reduces
with longer structures so that the B is again lowered for large T. This can be roughly expressed in terms
of a limiting T.B product of order 10^3 for inter-digital transducers and 10^4 for devices using grooved
reflective array structures which are not so widely available.

The dynamic range for a SAW device may be limited by the thermal noise floor or, more commonly, by spurious signals due to electrical breakthrough, triple transit, and bulk wave generation. The signal insertion loss enters into the equation and a dynamic range of 60 to 70 dB is typically attainable.

We now consider the range of processing functions readily available in SAW form, excluding those which are not yet past the research phase.

SAW Delay Lines

Fixed SAW delays of up to a few tens of μs are now finding their way into systems, for example to provide the 'early' and 'late' channels in a synchronization tracking loop in spread spectrum communications. In radars using a fixed prf of 10 kHz or more, SAW delays are convenient for clutter cancellers and coherent memory filters for doppler processing (Jackson, J.D. et al, 1976) where a wideband delay accurately matched to the pulse repetition interval is needed. The accurate registration can be maintained in spite of temperature variations by also using the delay line to generate the pulses. Delays variable in discrete steps are obtainable by switching between a set of alternative input or output transducers. The minimum delay increment possible is about 20 ns. Similarly, tapped delay lines use several output transducers simultaneously. Continuous control of the delay is possible using a pair of dispersive SAW delay lines (Maines, J.D and Paige, E.G.S., 1973; and Dolat, V.S and Williamson, R.C., 1976) as shown in Fig 1. In this scheme the dependence of delay upon frequency is exploited by mixing the signal to the frequency that corresponds to the required delay and subsequently translating back to the original centre frequency. Dispersion within the signal bandwidth is compensated by the second SAW disperser of opposite slope.

SAW Transversal Filters

In this important class of devices, the function of the inter-digital transducers is expanded to build into the device a response defined by the detailed positions and lengths of the electrodes. Within the previously mentioned limits of duration and bandwidth we can specify an arbitrary and "complex" impulse response, (complex since we can separately control both the amplitude and the phase of each contribution to the impulse response). The precision of the realised response is dependent on careful control of spurious secondary effects and it is a measure of the success achieved in this that errors are often at the -60 dB level.

The main successes of SAW devices have been based on variants of the transversal filter: those configured as IF bandpass filters, dispersive delay lines for waveform generation and matched filtering, and 'chirp' filters incorporated into fast-acting spectrum analysers. Filters with complicated amplitude and phase responses are now produced accurately and cheaply in production quantities for TV receivers whilst the ability to generate and filter optimally designed modulation waveforms has materially improved radar pulse compression systems. A dramatic picture of the possible processing bandwidth of a SAW transversal filter is provided by noting that for $T = 1$ μs, $B = 500$ MHz ($TB = 500$) a digital equivalent would require 10^3 products to be summed at a sampling rate of 1 GHz: a computing rate of 10^{12} products per second!

The principle of the compressive receiver (Edwards J A and Withers, M J., 1967) and its extension to a full Fourier transformer (Otto, O.W., 1976) greatly enlarges the scope for applying SAW chirp filters to pressing problems in signal analysis. These range from the rapid determination of frequency over wide bandwidths to adapting the frequency response of a filter using a versatile processor able to swap reversibly between the time and frequency domains, (Maines, J.D. et al, 1975).

The essential simplicity of the SAW Fourier transformer is indicated in Fig 2. A single chirp filter is used to compute the integral of products which arises in the definition of the transform. Pre- and post-multiplication (mixing) by chirp waveforms is also necessary, unless the amplitude (not phase) spectrum only is of interest, in which case the filter output is simply detected. The chirp signals are conveniently obtained by impulsing other SAW filters (not shown). Because SAW devices operate at IF, carrying complex (amplitude and phase) modulation, quadrature channels are not involved if the signal is also available at IF. The basis of the scheme is the identity

$$e^{-\mu\tau^2} \int (S(t)e^{-i\mu t^2}) \ . \ e^{i\mu(\tau-t)^2} \ dt \ = \ \int S(t)e^{-i2\mu\tau t} \ dt$$

The left hand side represents the hardware function{S is multiplied, convolved, multiplied("M-C-M") with functions showing quadratic phase or linear frequency with time} while the right hand side can be identified as the Fourier transform of S, if we interpret the output time scale, τ, as a frequency (f) axis scaled by $f = \mu\tau/\pi$. An alternative " CMC " arrangement is equally valid and time-weighting functions can be used to control spectral sidelobes down to at least - 35 dB. As would be expected, the spectral resolution is limited to ~1/T for the shortest chirp in the system.

SAW Processors with Variable Parameters

A frequent need in signal processing is the correlation or convolution of signals against a variable reference function, usually another signal. There is therefore a strong research interest in SAW convolvers and correlators. These exploit non-linear effects in the piezoelectric substrate, in a semiconductor or in an array of diodes (see Heighway, J. and Paige, E.G.S., 1976 for references to this work). However these methods of generating and summing product terms between signals involve either a high insertion loss, reduced bandwidth, limited number of correlation points or the fabrication of a narrow air-gap over the SAW substrate.

Possibly the most promising short term approach to the convolution and correlation of general signals using SAW is (as in digital operations), through the use of high speed Fourier transformation. (Gerard, H.M. and Otto, O.W., 1976). In general such completely general operations in SAW are still within the province of the research laboratory.

Variable parameter SAW processing is therefore at present largely restricted to the relatively simple cases of variable delay, variable bandpass filtering based on the chirp transform and a chirp filter of variable slope (Newton, C.O. and Paige, E.G.S., 1976). All these rely on ingenious application of fixed-parameter filters.

Agile Oscillators

Although perhaps on the fringe of signal processing proper, SAW oscillators deserve mention here particularly because of their programmability. Control of the frequency is possible both in continuous (~1%) FM mode and by controlling the excitation of discrete frequency modes. (Browning, I, et al, 1976). These techniques provide new modulation techniques and fast-acting low-noise synthesisers covering many tens of MHz.

Where the requirement is for fast frequency hopping, CW frequency modes can be selected by a SAW filter bank (Laker, KR et al 1976) or again a combination of chirp filters can be applied (Hannah, J.M. et al, 1976), to generate sum or difference tones by mixing time-offset chirp signals .

3.3   CCD Processing

There is a marked parallelism in the signal processing functions possible with CCD and SAW devices. This is because both are analogue delay lines within which the signal remains accessible during its passage through the line. The attractive features of CCD include:

     (i)   Low cost in volume production since IC technology is used.

    (ii)   'Digital' type flexibility by virtue of the clocked time scale.

   (iii)   High bandwidth (~10 MHz).

    (iv)   Long delays (many ms).

    (v)   Precise time control for system synchronization.

    (vi)   Small size and power consumption.

CCD offers almost uniquely (cf. the bucket brigade which it has almost completely superseded) the possibility of using digital concepts in the analogue domain. Through applying discrete time operations on sampled data, we have much of the adaptability of digital components without the hindrance of analogue-to-digital conversion and the reduced bandwidth of multi-bit arithmetic. The reverse side of the coin is that we suffer the drift sensitivity of analogue devices without gaining the unlimited accuracy and storage time characteristic of the computational approach. The key to success with CCD lies therefore in identifying those applications to which processing speed, variable clock rate, small size and power dissipation are important but to which extreme accuracy is not crucial.

In contrast with the fixed propagation velocity of SAW devices, the passage of date through a CCD is under user control, subject to a limiting storage time imposed by thermal liberation of charge carriers in the device, and subject to a maximum data rate made possible by the clock waveform generator. As with any sampled system, the sampling rate f_s (= the clock rate) places an upper limit of $\frac{f_s}{2}$ to the signal bandwidth so that a device with N storage locations can accommodate signals with a time-bandwidth product

$$T B \; = \; \frac{N}{f_s} \; . \; \frac{f_s}{2} \; = \; \frac{N}{2}$$

subject to the following qualifications:

     (i)   T cannot exceed the storage limitation noted above, which is often of order 100 ms.

    (ii)   B may be restricted by the cut-off transition of any anti-aliasing filter that precedes the CCD.

   (iii)   B may be further restricted by the smoothing of the signal within the CCD due to its finite efficiency of charge transfer. This preferentially attenuates the higher frequencies present in the signal with a response curve of the type shown in Fig3. (Vanstone, G.F. et al., 1974) for which a first order expression is

$$\text{Amplitude} \propto \exp \left[\; - N p \varepsilon \; (1 \, - \, \cos \, 2\pi f/f_s) \; \right]$$

             where p is the number of CCD electrodes per stored sample and ε is the fractional charge residual for each interelectrode transfer.

Maximum actual bandwidths, as distinct from those normalised to f_s, are governed by the capability of the clock generator. When this is on the CCD chip, power dissipation tends to limit B to about 10 MHz. External clocks are commonly used up to several times this but are of course less convenient in use.

The way in which the T and B capabilities of CCD and SAW complement one another is illustrated in Fig 4. Each has a comparable limit in T.B product, this being due in the CCD case to N being limited by achievable

charge transfer efficiencies, in the range 10^{-4} to 10^{-5}. Figure 3 cannot be regarded as definitive because of differences between device functions and between research and production achievements; however its intention is to depict the contemporary development stage for fixed-function transversal filters. For completeness a hypothetical convolver based on an existing single-chip digital multiplier capable of a 5 MHz data rate has been used to provide a comparison with digital technology. The useful dynamic range is determined by the non-linearity of input and output circuits and spurious contributions to the output due to clock waveform breakthrough and charge transfer residuals. 60 dB is often available.

To date, CCD technology has not matured to the stage where systems relying upon them are in regular production. Nevertheless a significant range of processing functions have been demonstrated and some of these can be expected to appear in systems very shortly.

CCD Delay Lines

Although precision delay may seem the obvious first application for CCD, and indeed several exploratory processors using them have been made, no firm requirement for it seems to have emerged. This is partly because of the established capability of high speed digital shift registers and RAM and partly through the cheapness of quartz delay lines, particularly for TV line delays. The dynamic range of a CCD is often inadequate for delay line cancellers in radar and the charge transfer efficiency is severely tested when a large number of range cells is involved, especially in recursive filters (Roberts, J.B.G. et al, 1973).

When advantage is taken of the freedom to change the clock rate however, the CCD delay line becomes far more competitive as a means of adapting the time scale of a signal. Speeding up recorded speech without changing the pitch, correction of speed variations during tape replay and gathering high speed transients for slow readout for analysis or recording all appear to be viable applications for CCD. In each of these cases obviating the need for an A-D convertor is an important advantage for CCD. Variable delays have also been used to explore the correlations present in signals as shown in Fig 5. (Roberts, J.B.G and Eames, R., 1975) and to dynamically focus ultrasonic arrays (Bernardi, R.B. et al, 1976).

Tapped CCD delay lines have been available for some time with 32 taps spaced at one-sample increments. This limit to the T.B product (almost inevitably forced by the package pin count) seems to be adequate for some users, particularly perhaps adaptive filtering with the tap weights driven externally, possibly using multiplying A-D converters driven by a microprocessor. However it may be that since the time-flexibility of the CCD is not exploited, and digital hardware is involved in any case, the CCD has no overwhelming advantage unless it proves possible to integrate the complete processing function on to a chip.

Such an application, in the transform coding of TV images, has been proposed (Roberts, J.B.G et al, 1977) in which a tapped CCD may be integrated with a matrix of resistors defining the transform.

CCD Transversal Filters

Functionally the transverse filter is identical to the tapped delay line. However using a _fixed_ set of tap weights removes the need for a separate pin connection per tap and allows a device with a T.B of several hundred to be incorporated on one chip. The summing of tap contributions to form the output is an integral function with the clocking and the very powerful computing capacity of such a relatively simply manufactured chip makes the transversal filter (as in the case of SAW) the foundation for the success of CCD technology in signal processing. To relate this to its digital equivalent, the filter reported by Buss et al (1976) with 500 points clocked at 500 kHz would require a multiplication rate of 250 MHz. This gives the CCD transversal filter a marked advantage over digital processing when simple, cheap filters are required for bandwidths up to a few MHz (see Fig 4).

Superficially there is a strong resemblance between the structure of SAW and CCD transversal filters since both define the impulse response by the positioning of breaks in the fingers of an interdigital structure. However the geometrical scale and the physical effects used are completely different. An important degree of freedom is exclusive to the CCD filter, that of adjusting the time (and consequently the frequency) scale by simply changing the clock rate. Such an adaptive tunable filter may be the first CCD to find a volume market, in citizen's band radio receivers ('Electronics', 1977).

The performance limitations for CCD filters are determined by fabrication accuracy, the noise level of the output amplifier, the linearity of the input and output circuits and to some extent by the charge transfer inefficiency. A dynamic range of 75 to 80 dB has been reported for a 500-point filter with 1% harmonic distortion (Buss, D.D. et al, 1976).

CCD matched filters for chirp and phase-coded sequences have performed well for T.B products (=processing gain) of order 100 and bandwidths of a few MHz. Frequency filters with 40 dB stopband rejection are also possible with Q factors greater than 100.

As with SAW technology, one of the most exciting prospects held out by the CCD transversal filter capability is that of using 'chirp-transform' methods for spectrum analysis and Fourier transformation. In the lower B and longer T regimes appropriate to CCD an analyser can be constructed very compactly by combining several of the chirp filters needed on one chip. This can be carried further to provide for instance a multi-channel analyser capable of processing velocity data from many ranges in a radar system (Buss, D.D et al, 1975).

Because CCD is a time-discrete technology, we obtain the discrete Fourier transform (DFT) or Z-transform characteristic of sampled data (hence the name 'chirp-Z-transform' or CZT). Other variations are possible, for instance the discrete cosine transform, the 'sliding' DFT (Broderson, R.W. et al, 1976) and the prime transform which evaluates the DFT without using multipliers (Rader, C.M. 1968 and Buss, D.D. et al, 1976B) CCD, like digital processors are normally operated at base band (zero centre frequency) so that the real and imaginary parts of complex data must be handled separately, for instance four convolvers being required for a chirp-Z Fourier transformer, as indicated in Fig 6. In view of the fact that several filters can be

incorporated on a chip, this is not a big drawback. A spectrum analyser with 500 resolved frequencies and -55 dB dynamic range has been reported showing charge transfer residuals -35 dB below the spectral peaks (Buss, D.D. et al, 1976A).

Other CCD Processors

Three further CCD functions are of considerable interest for signal processing: de-multiplexing, time-delay-and-integration (TDI) and correlation, although the last of these is a largely unattained objective.

De-multiplexing, or splitting serial data into parallel channels, is in itself not a major accomplishment; but using both dimensions of the silicon surface allows a CCD processor to include within it this function shown schematically in Fig 7. This has been applied to sorting radar data by range in doppler processors (Buss, D.D. et al, 1975 and Roberts, J.B.G et al, 1976).

The TDI is almost the inverse of the de-multiplexer in that it accepts multiple inputs in parallel and produces a single output. Rather than merely serialising a data stream, it adds together samples by pooling the corresponding charge packets so that the output is a time-staggered sum of the multiple inputs as sketched in Fig 8. This function is required in mechanically scanned infra red imagers and may be integrated with the imager (Nummedahl, K. et al, 1975). The CCD clock rate is readily synchronised with the system scan. A similar addition of time staggered signals is appropriate in array beamforming and since CCD scales of time and bandwidth are appropriate to sonar systems and ultrasonic scanners this appears an attractive method of beam steering since a change of clock rate automatically adjusts all the delay increments. (Melen, R.D, et al, 1975).

A completely general correlator is the holy grail of the signal processor because so many other operations (filtering, pattern recognition, Fourier transformation, etc) can be expressed as a correlation of a signal against an appropriate reference. The cross-correlation between a function x(t) and a reference y(t) (which may be x itself in which case we obtain the auto-correlation function) is the function ϕ_{xy} of time displacement, τ,

$$\phi_{xy}(\tau) = \int x(t) \; y \; (t-\tau) \; dt$$

where the integration is over the available range of t (ideally infinite). In CCD systems the integral is replaced by a sum of products of discrete samples. A correlator has therefore to introduce time lags τ corresponding to the argument of ϕ_{xy} required, to form products and integrate these.

We have seen the possibility of a limited form of correlator (Fig 5) in connection with CCD delay lines. However this wastes most of the available information and is appropriate only to stationary signals because ϕ is measured for only one value of τ at a time. A more efficient correlator can be built around a tapped line providing several values of τ simultaneously and using a corresponding number of analogue multiplier (Fig 9). A further possibility is the correlator using two tapped CCD lines shown in Fig 10. This provides a programmable filter, the impulse response or correlation reference being read into one register and held (up to the storage time limit, typically 30 ms) while the signal is clocked through the other. This scheme has been successfully implemented on a single chip with 4-quadrant multipliers allowing 40 dB output dynamic range (Bosshart, P., 1976).

The most successful CCD correlators to date have operated in a hybrid analogue/digital fashion in which analogue samples in a tapped CCD are connected, or not, to summing lines; the connections being determined by the individual bits of a digital reference function. With appropriately weighted summing (using coefficients of the type 2^n) an analogue correlation output is obtained. It is not clear that this approach has a significant advantage over all-digital realisations even though arithmetic multiplications are avoided. Perhaps its most appropriate application is in correlating signals containing binary codes since in this case an analogue signal and one-bit reference are needed (Herrmann, E.D. et al, 1976).

It seems likely that the most useful CCD correlator will be based on Fourier transformation. When this technique is fully developed in a convenient form, correlation and convolution will be reduced to multiplication in the transform variable domain (usually frequency), followed by an inverse transform. Meantime it cannot yet be claimed that CCD technology yet offers a simple or universally convenient means of correlating signals.

3.4   Inter-relation between Digital, SAW and CCD Technologies

We have reviewed the present capability for signal processing offered by three important technologies. In general we find marked differences between them highlighted by their different time delay, T, and bandwidth, B, parameters, illustrated in Fig 4; but also due to the characteristic adaptability of digital components which contrasts almost completely with the invarient properties of most SAW devices. CCD occupies an intermediate position in many respects including T, B and flexibility. Both SAW and CCD justify their commercial survival by their high bandwidth filtering capability (CCD also has important roles in digital stores and imagers which we have not covered here), the CCD having an important degree of freedom in its adjustable time scale which the SAW cannot match.

We have seen however, several ways in which fixed-response filters can be used to form more general processors such as variable delays, variable slope chirp filters, Fourier transformers and hence correlators and convolvers. The range of processing functions possible with SAW and CCD is essentially similar but their complementarity in parameter values suits them for completely different uses, as is summarised in Tables 2 and 3.

The overlap between digital and CCD capabilities evident in Fig 4 represents an area which digital technology can be expected to make largely its own whilst the SAW-CCD overlap may be seen as an advantage to both since it permits exchanging signals between the two technologies.

TABLE 2 - Typical Device Parameters

	SAW	CCD
Storage time T	≤ 100 μs	$\leq$ 1s
Bandwidth B	< 600 MHz	< 10 MHz
Flexibility	Essentially fixed devices	Discontinuous operation, 4-5 decades of bandwidth possible with one device
TB	$< 10^3$ common 10^4 difficult	100-400 filters 10^5 image storage
Centre frequency	20-200 MHz	Usually zero
Useful Dynamic Range	70 dB	70 dB
Spurious effects	Triple transit Bulk waves Reflections	Charge transfer inefficiency Clock feed through Non-linear input
Stability	1 in 10^6 short term long term stability not as good	Extremely sensitive to circuit conditions
Device Size	A few cm^2 typically	IC package

TABLE 3 - Areas of Signal Processing Application

	SAW	CCD
Instrumentation	Spectrum Analysis Frequency Synthesis Pressure Temperature } Sensors	Spectrum Analysis Signal averaging Transient Recording Time compression of speech
Communications	Oscillators IF filters, Wideband Rapid Freq Measurement, spread Spectrum modems	Channel equalisation TV Image Coding Speech processing Scrambling High resolution frequency measurement
Radar	Single pulse processing: dispersion/compression, STRETCH, beam forming, pseudo clutter reference	multi-pulse processing: MTI Doppler Integration
Sonar		Beam forming. Beam scanning
Infra Red		Integration of scanned images
Ultrasonics		Beam steering, focusing

these considerations raise a more constructive argument: given the existence of the different technologies how best can we exploit them together? This matter has received some attention recently (Beauvais, T et al 1976; Roberts, J.B.G., 1976; Whitehouse, H.J. and Speiser, J.M., 1976) and offers promise of combining the best features of each, for example the adaptability and long storage times of digital operation with the compactness and low power consumption of fast SAW processing. We therefore conclude this review with some examples of combined-technology processors which hold promise for expanding the power of signal processing subsystems.

Many systems involve the processing of many data channels of individually low bandwidth. For instance a passive sonar may use 100 'beams' or an active radar or sonar, a large number of range-azimuth cells, for each of which spectrum analysis is required covering only a few kHz with a resolution measured in Hz or tens of Hz. SAW time scales are completely inappropriate to the data as it stands but by storing the data for time-compressed readout and spectrum analysis, we can at once match the data to the fast processor and accomplish the analysis in a short enough time for the single analyser to be applied sequentially to all the parallel channels in real time. Beauvais et al, (1976) have described such a passive sonar system with a digital store giving a compression ratio of 3.10^4 while Roberts et (1977A) has described radar doppler processors using both digital and CCD memories with time compression factors of order 10^3. The attraction of the CCD store is that a compact simple system without A-D converters is possible by using special CCD structures (Roberts, J.B.G, et al, 1976) while a digital store has a longer storage time, can be read non-destructively and uses standard components only. Either storage medium allows the same hardware modules to accommodate a wide range of system parameters by using clock waveforms appropriate to the system pulse rate and bandwidth. The same principle has been suggested as attractive for miniaturising the processing of synthetic aperture signals in sideways-looking airborne radar. In this case correlation against a chirp reference, rather than spectrum analysis, is required to which a SAW filter is equally well adapted.

The mere fact of adjusting the data to the fixed SAW processor is an important extension to SAW technology, enabling it to access perhaps 5 more decades in bandwidth than would otherwise be possible. Even without exploiting the multi-channel capability we can for instance use a SAW chirp analyser as a highly flexible spectrum analyser with a constant TB, or number of resolvable frequencies, over a wide range of bandwidths. The effective dynamic range as specified by the highest spectral sidelobe is presently only 35 to 40 dB but the hardware simplicity should make the cost very competitive with other types of analyser.

The virtues of combining CCD, digital and SAW technologies in a system have been pointed out by Roberts(1976) in connection with a spread spectrum video link. CCD technology (in addition to its role as the imager) can be used for coding TV video to remove redundancy either using chirp techniques for Fourier related transforms (Wrench, E.R. et al, 1976) or using a tapped CCD delay line to realise other transforms (Roberts, J.B.G. et al, 1977B). A further reduction in redundancy is made using differential pulse code modulation (DPCM) coding in which advantage of its digital nature can be taken to make the coding parameters programmable. When the message redundancy has been reduced as far as possible, the spectrum of the transmitted signal is spread using a combination of SAW chirp filters to generate a sequence of fast frequency hops (Hannah, J.M, et al, 1976) covering perhaps 100 MHz.

Profitable combinations of digital with CCD techniques seem possible in two distinct ways. Firstly we can control and adapt CCD filtering using logic (Herrmann, E.P., et al 1976 and Dilley, J.E. et al, 1976), the simplest adjustment being to the clock rate, for tuning a bandpass filter (Electronics, 1977). Secondly we can apply CCD technology in a digital environment (Zimmerman, T.A. and Allen, R.A., 1976) with projected advantages in power consumption and high functional density. Other avenues in this direction are offered by the possibilities for using CCD's to carry multi-level digital data. Ideas for applying residue arithmetic and multipliers that operate by convolving binary coefficients have been advanced as a means towards fast digital correlators and convolvers based on CCD techniques (Lagnado, I, et al, 1976). However it seems too soon to evaluate the competitiveness of these against more conventional digital processors.

4   CONCLUSION

The advent of the SAW and CCD technologies cannot be seen as a threat to the continuing development of digital technology and techniques. However their attributes are sufficiently distinct from each other and from digital LSI to ensure a place for them in military and civil systems of the future. The T-B area of contention between CCD and digital devices is certain to be resolved in favour of the components backed by the requirements of the computer industry. However both CCD and SAW transversal filters have a combination of convenience and processing bandwidth well ahead of their digital equivalents and since these can be applied to Fourier transformation, and hence to a wide range of variable filters, there seems an assured future for SAW and CCD. The likelihood is that, as in the case of SAW, specialised CCD structures will be marketed as an integral part of a system with only a small range of general-purpose delay lines, tapped delays and filters becoming available generally.

The value of having each of these technologies available together is already beginning to be exploited and this trend can be expected to continue.

5   REFERENCES

Beauvais, T. et al, 1976 'Interfacing and Comparison of SAW and Digital Technologies', IEE Conf. Proc. No 144, pp 44-52.

Bernardi, R.B. et, 1976, 'A dynamically focused annular array', Proc IEE Ultrasonics Symp. pp 157-159.

Bosshart, P., 1976, 'An Integrated Analog Correlator using CCD', IEEE Int. Solid State Circuits Conf. Digest, pp 198-199.

Broderson, R.W. et al, 1976, 'A 500-Stage CCD Transversal Filter for Spectral Analysis', IEEE J. Solid-State Circuits, SC-11, pp 75-84.

Browning, I. et al, 1975, 'A SAW frequency synthesiser', Proc. IEEE Ultrasonics Symposium', pp 245-247.

Buss, D.D. et al, 1975, 'Radar Video Processing using the CCD Chirp Z Transform', CCD'1975 Proc, pp 283-290.

Buss, D.D. et al, 1976A, 'Charge Coupled Devices for analogue signal processing'. IEE Conf. Proc. No 144, pp 17-26.

Buss, D.D. et al, 1976 B, 'Spectral Analysis using CCDS', CCD '76 Proc. pp 208-218.

Dilley, J.E. et al, 1976, 'A Microcomputer-Controlled Adaptive CCD Transversal Filter', CCD '76 Proc, pp 269-276.

Dolat, V.S. and Williamson, R.C., 1976, 'A Continuously variable delay line system', Proc. IEEE. Ultrasonics Symposium, pp 419-423.

Edwards, J.A. and Withers, M.J., 1967, 'High Speed Spectrum Analyser using a pulse compression technique', Proc. IEE, 114 (11) pp 1613-1616.

'Electronics' 1977, 'TI introduces microprocessor-based citizens' band radios', May 12, pp 31-32.

Gerard, H.M. and Otto, O.W., 1976, 'A Programmable Radar Pulse compression filter utilizing Chirp Transform Correlation', Proc. IEEE Ultrasonics Symp, pp 371-375.

Hannah, J.M. et al, 1976, 'Fast Coherent Frequency Hopped Waveform Synthesis using SAW devices', Proc IEEE Ultrasonics Symp, pp 428-431.

Heighway, J. and Paige, E.G.S., 1976, 'Analogue Signal Processing with SAW devices', IEE Conf. Publication No 144, pp 27-43.

Herrmann, E.P., 1976 'CCD Programmable Correlator', CCD '76 Proc. pp 232-237.

IEE Conference Publication No 144: 'The Impact of New Technologies in Signal Processing' Aviemore 1976.

Jackson, J.D. et al, 1976, 'Pulse Doppler Processing with a SAW delay line', Proc. IEEE Ultrasonics Symposium, Annapolis, pp 505-509.

Lagnado, I, et al, 1976, 'A CCD Analog and Digital Correlator', CCD '76 Proc, pp 238-258.

Laker, K.R. et al, 1976, 'A Circuit approach to SAW filter banks for Frequency Synthesis', Proc. IEEE 64, pp 692-695.

Maines, J.D. and Paige, E.G.S., 1973, 'Surface Acoustic Wave Components, Devices and Applications', IEE Reviews, 120, pp 1078-1110.

Maines, J.D. et al, 1975, 'A novel SAW variable frequency filter', Proc. IEEE Ultrasonics Symp, pp 355-358.

Melen, R.D. et al, 1975, 'CCD Dynamically Focussed Lenses for Ultrasonic Imaging Systems', CCD '75 Proc, pp 165-171.

Newton, C.O. and Paige, E.G.S., 1976, 'SAW dispersive filter with variable linear frequency-time slope'. Proc. IEEE Ultrasonics Symp, pp 424-427.

Nummendahl, K. et al, 1975, 'Extrinsic Silicon Monolithic Focal Plane Array Technology and Applications', CCD '75 Proc. pp 19-30.

Otto, O.W., 1976, 'The Chirp Transform Signal Processor', Proc. IEEE Ultrasonics Symposium, pp 365-370.

Rader, C.M., 1968, 'Discrete Fourier Transforms when the number of Data Samples is Prime', Proc. IEEE 56 pp 1107-1108.

Roberts, J.B.G. et al, 1973, 'Moving Target Indicator Recursive Radar Filter using Bucket Brigade Circuits', Elect. Lett 9 No 4, pp 89-90.

Roberts, J.B.G. and Eames, R., 1975, 'A swept delay Correlator', CCD '75 Proc., pp 319-321.

Roberts, J.B.G. et al, 1976, 'A Processor for Pulse-Doppler Radar', IEEE J. Solid State Circuits SC-11 No 1, pp 100-104.

Roberts, J.B.G., 1976, 'CCD and SAW: Friends or Foes?', IEE Conf. Proc. No 144, pp 53-60.

Roberts, J.B.G. et al, 1977A, 'A New Approach to Pulse Doppler Processing', Proc. IEE Conf.'Radar -77,' London, Oct 1977, to be published.

Roberts, J.B.G. et al, 1977B, 'Transform Coding using CCD', Elec. Lett. 13 No 10, pp 277-278.

Vanstone, G.F. et al, 1974, 'The Measurement of the charge residual for CCD transfer using impulse and frequency responses', Solid-State Electronics, 17, pp 889-895.

Whitehouse, H.J. and Speiser, J.M., 1976, 'Signal Processing with combined CCD and Digital Techniques', IEE Conf. Proc. No 144, pp 61-65.

Wrench, E.R. et al, 1976, 'Analog Transforms:Practical Considerations', IEE Conf. Proc. No 144, pp 104-111.

Zimmerman, T.A. and Allen, R.A., 1976,'CCD Digital Arithmetic Functions:Experimental Results,'CCD '76 Proc, pp 190-196.
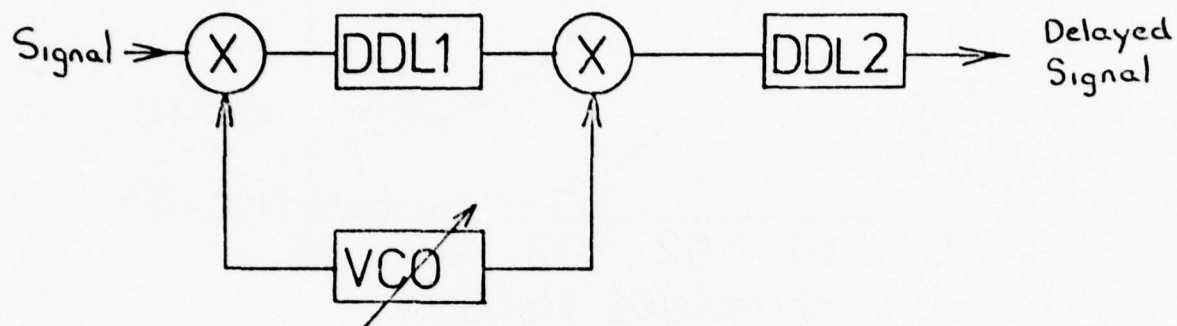
Fig.1   Continuously variable SAW delay line

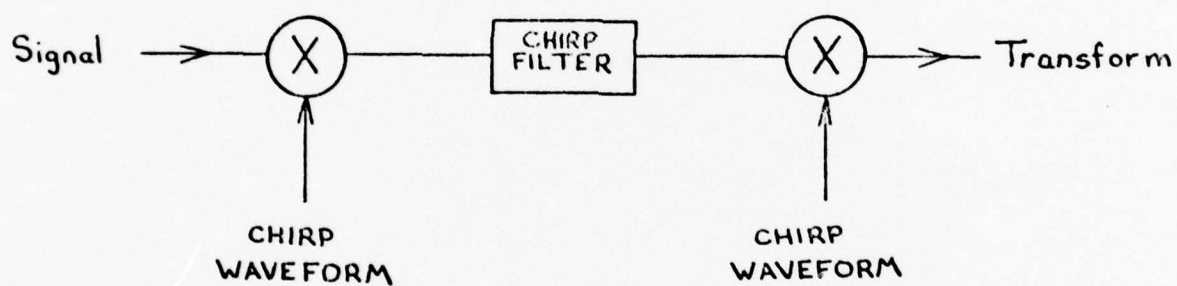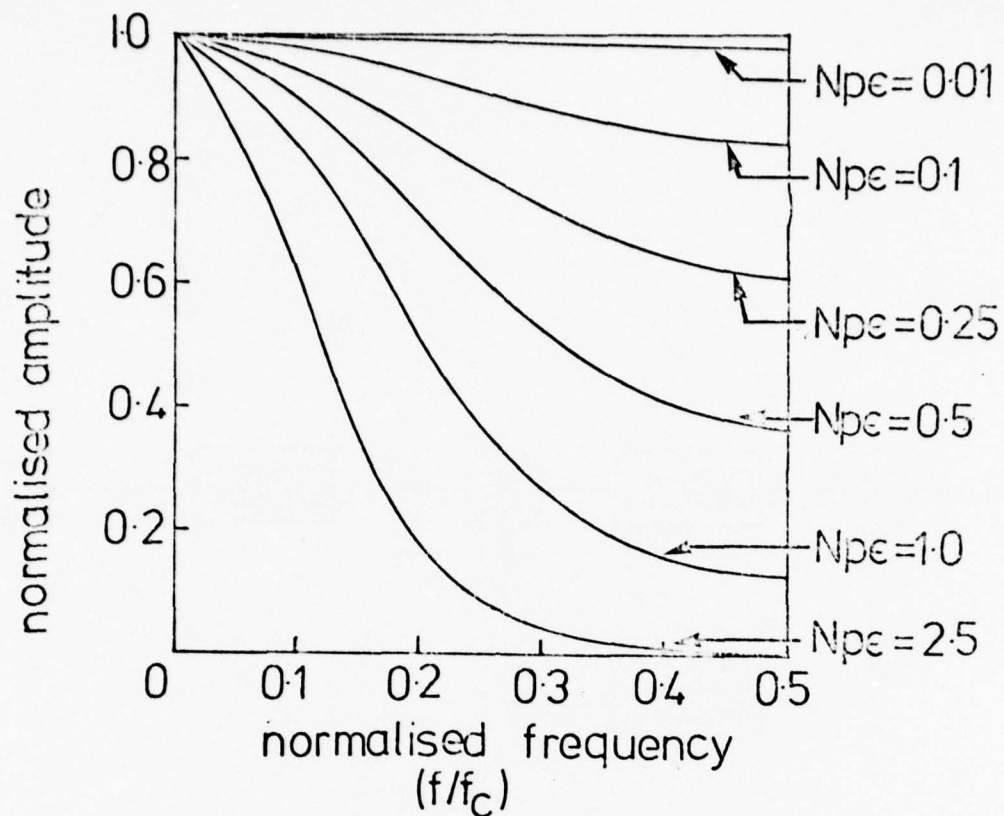Fig.2   SAW chirp transform schematic

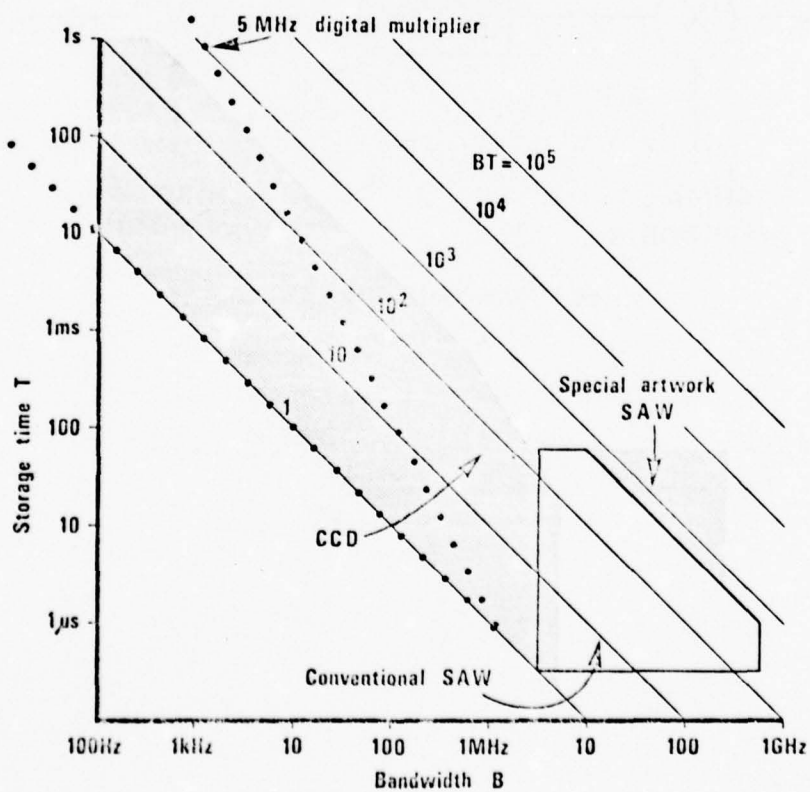Fig.3 CCD bandpass response due to charge transfer inefficiency

Fig.4 Time-bandwidth regimes for SAW, CCD and digital technologies
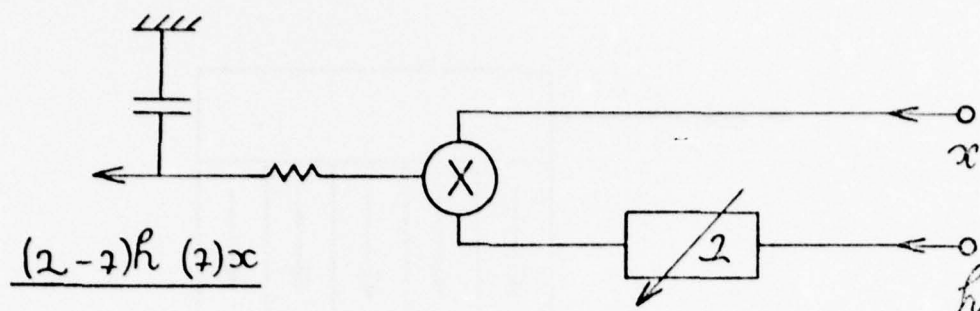
Fig.5 CCD swept delay correlator
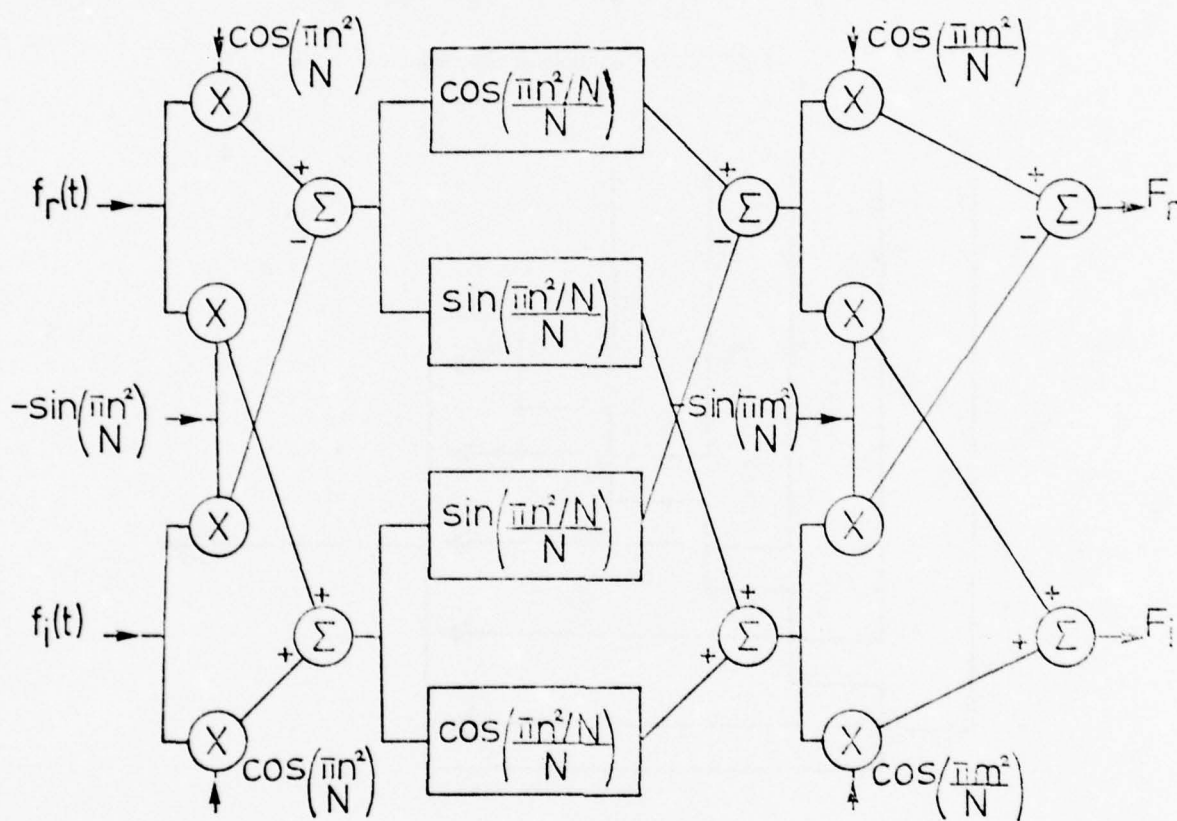
Fig.6 CCD chirp-Z transform schematic
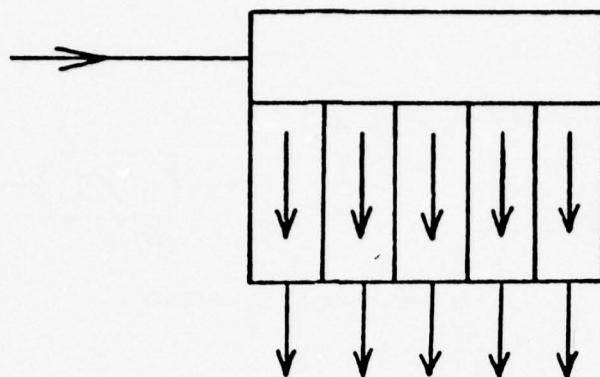
Fig.7   CCD demultiplexer schematic
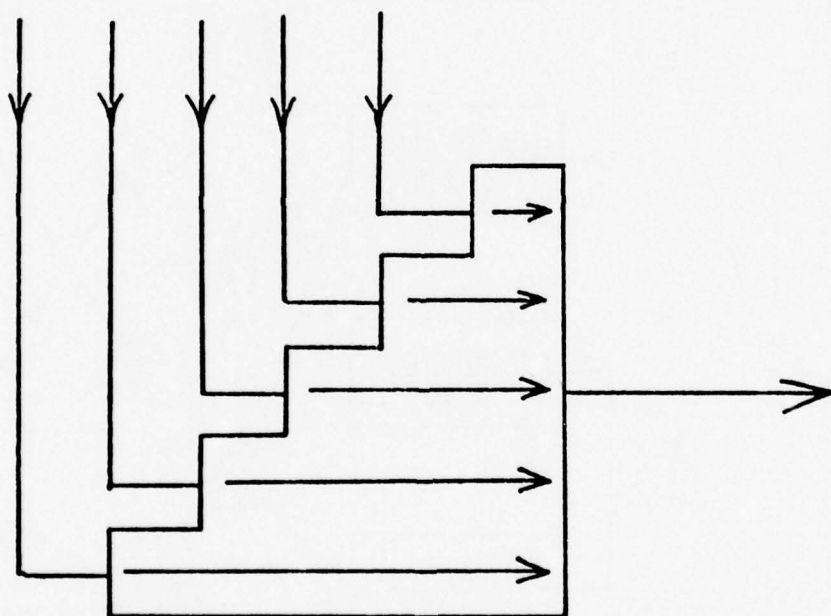
Fig.8   CCD TDI processor
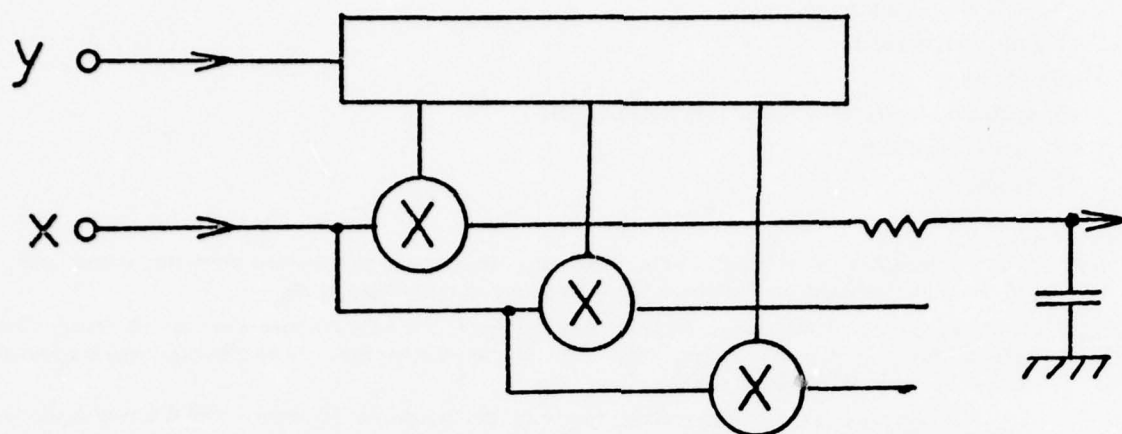
Fig.9   CCD tapped delay correlator
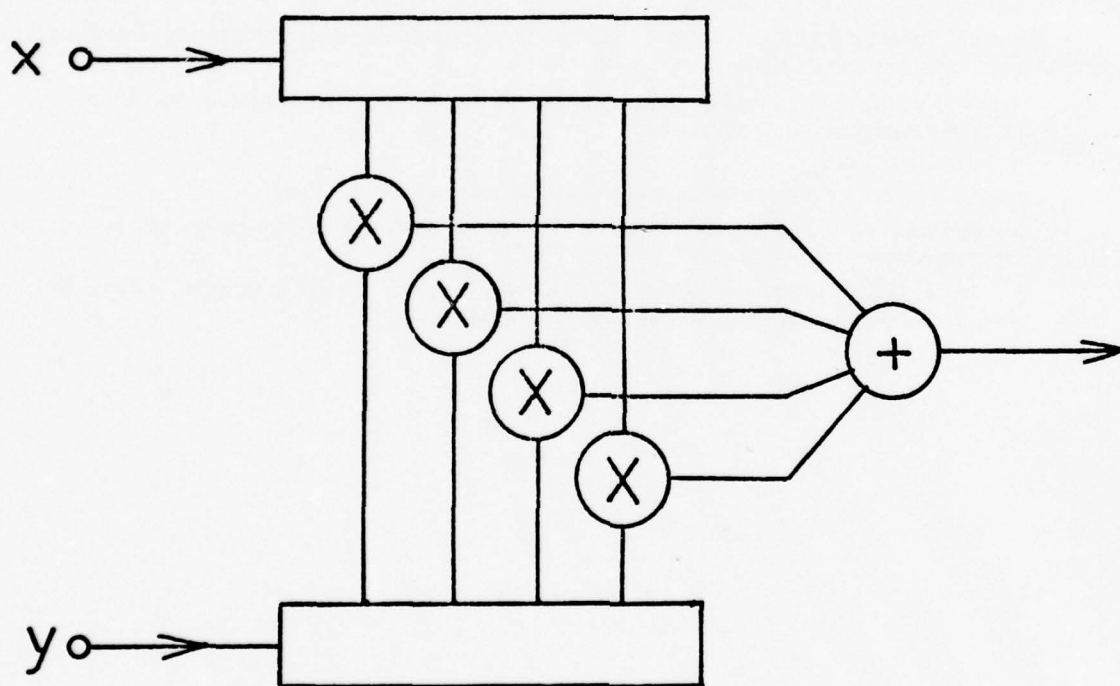
Fig.10   CCD stored reference correlator

DISCUSSION

E.Schaeffer

In the two techniques, SAW and CCD, can you comment on the following relative characteristics?

1 — Noise figure, noise temperature

2 — Input/output impedance

3 — Signal scaling

4 — Susceptibility to electric/magnetic field regarding EMI

5 — Nuclear susceptibility

6 — Reliability

Author's Reply

SAW and CCD technologies are very different both because of the physical principles on which they depend and because of their differing stages of development. I will summarize the two separately.

SAW: 1 — Noise figure is perhaps an inappropriate figure of merit. The SAW is a passive device with insertion loss varying between 2 db (band pass filter) and 70 db (non-linear convolver). 70 db dynamic range is achievable, limited by spurious bulk mode signals.

2 — Devices are normally tuned to present impedances in the region of 50 ohms — VSWR is very application-dependent: 1.2 over 20% bandwidth is perhaps typical.

3 — Maximum input voltages about 200 volts.

4 and 5 — I have no quantitative results to quote but SAW are perhaps amongst the most immune to EMI. They are normally completely sealed in a metal package and are specified for UK operational military systems, including airborne ones.

6 — Reliability is very high, probably the most sensitive feature being the metallic bonds to the transducer pads common to nearly all electronic devices.

CCD: 1 — CCD is a low noise technology, the output amplifier normally providing the noise floor 60 to 70 db below the maximum output signal level.

2 — Typical values:
Input impedance 10^9 ohms in parallel with 5 p F. Output impedance 1 Kohms.

3 — Maximum signal 2 volts peak-to-peak is typical. An on-chip amplifier is normally aimed for unity gain, otherwise output is in mV region.

4, 5 and 6 — I don't believe that CCD's have specifically been characterized in the these respects. However, they may be expected to behave similarly to other MOS devices.

MATERIAL CHOICE FOR OPTIMUM SAW DEVICE PERFORMANCE

Robert M. O'Connell, Andrew J. Slobodnik, Jr., and Paul H. Carr
Deputy for Electronic Technology (RADC/AFSC)
Hanscom AFB, Massachusetts 01731

ABSTRACT

The theory of surface acoustic wave propagation is reviewed and some of the various material design parameters which follow from that theory and which must be considered in making the optimum SAW device substrate choice are discussed. The parameters covered include SAW velocity, piezoelectric coupling constant, electromechanical power flow angle, temperature sensitivity, propagation losses, and beam steering and diffraction. Depending upon the device being designed and the application, some of these parameters are more important than others. In the design of temperature stable, broadband, low insertion loss devices, the important requirements are a zero temperature coefficient of time delay and a large piezoelectric coupling constant. Alternatively, the design of high frequency devices requires low loss substrate materials with large SAW velocities. The state-of-the-art in the development of new materials for these two classes of devices is reviewed.

1. INTRODUCTION

Because of their small size and low cost, surface acoustic wave (SAW) devices are becoming increasingly attractive for many applications (IEEE Proceedings, 1976). Fundamental to the optimum design of a SAW device is the proper choice of substrate material. A wise decision concerning this choice requires an understanding of the various material factors and tradeoffs inherent in the design of a SAW device. These elements will be discussed in an attempt to provide useful guidelines for the design of SAW devices up to microwave frequencies. The paper will begin with a brief outline of the theory of SAW propagation and the four important factors which describe it. These include SAW velocity, piezoelectric coupling constant, electromechanical power flow angle, and temperature coefficient of time delay. This will be followed by a discussion of various sources of material-related insertion loss, including propagation losses and beam steering and diffraction. The manner in which these concepts are applied depends upon the specific application at hand. To demonstrate this, the remainder of the paper will describe the application of these guidelines in materials related efforts to improve two classes of SAW devices: (1) temperature stable, broad-band, low insertion loss devices, and (2) high frequency devices.

One of the major sources of overall device insertion loss at microwave frequencies is propagation loss or attenuation. Not only is the magnitude of this phenomenon important for predicting absolute insertion loss and dynamic range, but its frequency dependence is equally important when, for example, designing filters having particular bandpass characteristics. The experimentally determined magnitude and frequency dependence of loss for several important surface wave materials will be presented.

The combined effects of beam steering and diffraction are also important. Diffraction causes the acoustic beam to lose its rectangular shape while beam steering results in the beam being offset from the desired location (the output transducer). Both contribute to device insertion loss in a manner which is not simply the sum of the separate constitutents. Thus, in the general case, design curves must be more specific than the universal information possible when each loss mechanism is considered individually. A quantitative measure of the extent of both beam steering and diffraction is provided by the slope of the power flow angle. A value of zero for this parameter implies isotropic diffraction and no beam steering, while a value of - 1 implies minimum diffraction with increased beam steering. These ideas will be discussed in detail.

The design of a broad-band, low insertion loss SAW device with temperature independent performance characteristics requires a substrate material having large piezoelectric coupling and zero temperature coefficient of time delay. Presently, the most readily available SAW substrate materials are lithium niobate and quartz. As can be seen in Fig. 1, lithium niobate has very large piezoelectric coupling ($\Delta v/v$), which is essential for broad-bandwidth and low insertion loss. However, it also has a large temperature coefficient of time delay, which necessitates the use of ovens and other schemes for temperature control. The ST-cut of quartz (Schulz, M.B., ... 1970), on the other hand, is temperature compensated, but it has only 1/40th the piezoelectric coupling of lithium niobate, a definite disadvantage for low insertion loss devices. Consequently, recent developments in the search for temperature compensated materials having piezoelectric coupling greater than that of quartz are of interest and will be discussed.

Currently the upper frequency limit of optically fabricated SAW devices operating at a fundamental mode is about 1 GHz. The design of higher frequency devices requires either electron beam fabrication, advances in the state-of-the-art of optical lithography (Budreau, A.J., ... 1975), or new substrate materials whose SAW velocities are larger than those of either quartz or lithium niobate, both of which are about 3000 m/sec. Certain materials like aluminum nitride, with SAW velocities of about 6000 m/sec, can extend the frequency limit of SAW devices to 2 GHz. Progress in obtaining reproducible thin films of aluminum nitride will be summarized.

2. SURFACE ACOUSTIC WAVE PROPERTIES

2.1 SAW Propagation Theory

The concept of the generation and propagation of a surface acoustic or Rayleigh wave (Rayleigh, Lord, 1885) is depicted in Fig. 2. The electromagnetic signal is converted to an acoustic wave at the input interdigital transducer (IDT) (White, R.M., ... 1965) by means of the piezoelectric effect. The delayed acoustic wave is then converted back to an electromagnetic signal at the output IDT. The energy of the SAW decays exponentially into the material and is generally confined to within a few acoustic wavelengths of the surface.

A quantitative description of this process requires solution of the problem of acoustic wave propagation on an arbitrary anisotropic piezoelectric medium. That problem has been solved (Campbell, J.J., ... 1968) (Jones, W.R., ... 1969), and an outline of the solution follows. The coordinate system used to define the problem is shown in Fig. 3. The thin conducting plane shown is used to calculate an estimate of the piezoelectric coupling constant, as will be discussed further in the next section.

The solution to the problem is obtained by solving the continuum equations of motion together with Maxwell's equations under the quasi-static assumption, the strain-mechanical displacement relations, the piezoelectric constitutive relations, and the appropriate boundary conditions. When the linear equations describing these quantities (Tiersten, H.F., 1963) (Slobodnik, A.J., Jr., 1976) are combined, the following differential equations for the components of mechanical displacement u_i and the electric potential φ within the crystalline medium are obtained:

$$c'_{ijkl}u_{k,li} + e'_{kij}\varphi_{,ki} = \rho\ddot{u}_j, \qquad j = 1, 2, 3 \left.\begin{array}{c} \\ \\ \end{array}\right\} \quad x_3 > 0 \tag{1}$$

$$e'_{ikl}u_{k,li} - \epsilon'_{ik}\varphi_{,ki} = 0 \tag{2}$$

Outside the crystalline medium, Laplace's equation describes the electric potential.

$$\nabla^2\varphi = 0. \qquad\qquad -h \le x_3 \le 0 \tag{3}$$

In equations (1) and (2), the primed quantities are the elastic constants (c'_{ijkl}), the piezoelectric constants (e'_{ijk}), and the dielectric constants (ϵ'_{ij}) in a rotated coordinate system obtained through the Euler transformation matrix (Goldstein, 1950) in which wave propagation is always along the 1 direction. Note that the summation convention (over 1, 2, 3 for repeated indices is employed. Also, the dot notation refers to differentiation with respect to time, while an index preceded by a comma denotes differentiation with respect to a space coordinate.

Solutions of equations (1) and (2) are assumed to be of the standard complex traveling-wave form in which v_s is the wave velocity, α the exponential decay into the crystal, and ω the steady-state angular frequency.

$$u_i = \beta_i \exp[-\alpha\omega x_3/v_s] \exp[j\omega(t - x_1/v_s)]. \qquad i = 1, 2, 3 \tag{4}$$

$$\varphi = \beta_4 \exp[-\alpha\omega x_3/v_s] \exp[j\omega(t - x_1/v_s)]. \tag{5}$$

The displacements and potentials are considered to be independent of the x_2 coordinate.

Substituting (4) and (5) into (1) and (2) yields a linear homogeneous system of four equations in the unknowns β_1, β_2, β_3 and β_4. The determinant of the coefficients of the unknowns in these equations must be zero in order that a nontrivial solution exist, i.e.,

$$A_8\alpha^8 + jA_7\alpha^7 + A_6\alpha^6 + jA_5\alpha^5 + A_4\alpha^4 + jA_3\alpha^3 + A_2\alpha^2 + jA_1\alpha + A_0 = 0 \tag{6}$$

where the coefficients A_n, $n = 0, 1, \cdots, 8$, are purely real and a particular value of v_s has been assumed. Since the fields must be bounded, or go to zero as $x_3 \to \infty$, only the roots with nonnegative real parts are allowed. In addition, these roots are either pure imaginary or occur in pairs with positive and negative real parts. In general, roots occur such that four (three for nonpiezoelectric crystals) with positive real parts can be selected. However, if for a given velocity four such roots cannot be found the possibility of degenerate surface waves must be pursued. Upon obtaining the admissible values of α from equation (6), corresponding values of β_i (to within a constant factor) can be found for each α from the linear homogeneous system cited above.

The total fields (mechanical displacement and potential) may now be expressed as a linear combination of the fields associated with the admissible values of α. For $x_3 > 0$

$$u_i = \sum_{l=1}^{4} B^{(l)}\beta_i^{(l)} \exp[-\alpha^{(l)}\omega x_3/v_s] \exp[j\omega(t - x_1/v_s)]. \quad i = 1, 2, 3 \tag{7}$$

$$\varphi = \sum_{l=1}^{4} B^{(l)}\beta_4^{(l)} \exp[-\alpha^{(l)}\omega x_3/v_s] \exp[j\omega(t - x_1/v_s)]. \tag{8}$$

In the region $-h \le x_3 \le 0$, the potential is a solution of Laplace's equation (3). A solution satisfying the continuity condition at $x_3 = 0$ and vanishing at $x_3 = -h$ is

$$\varphi = \sum_{l=1}^{4} B^{(l)}\beta_4^{(l)} \operatorname{csch}\left(\frac{\omega h}{v_s}\right) \sinh\left(\frac{\omega}{v_s}(x_3 + h)\right)\exp[j\omega(t - x_1/v_s)]. \quad -h \le x_3 \le 0. \tag{9}$$

Mechanical and electrical boundary conditions (Campbell, J.J., ... 1968) (Jones, W.R., ... 1969) must also be satisfied by substituting the waveforms of equations (7) to (9) into the appropriate expressions for these conditions. This yields a set of homogenous equations for the so-called partial field amplitudes $B^{(l)}$. The transcendental equation obtained by setting the determinant of the matrix of coefficients of this system equal to zero determines the surface-wave velocities for a given set of $\alpha^{(l)}$.

Once equation (6) and the set of $\beta^{(l)}$ equations have been simultaneously solved by computer iterative techniques (Jones, W.R., ... 1969) for the actual set of $\alpha^{(l)}$ (with associated $\beta_i^{(l)}$) and the actual surface wave velocity, the partial field amplitudes $B^{(l)}$, $l = 1, 2, 3, 4$ may be calculated to within a constant factor.

These amplitudes are used directly to evaluate the components of the mechanical displacement of equation (7) and the electric potential of equation (8). The components of the electric field, strain, electric displacement, and stress as functions of ωx_3 follow from Maxwell's equations, the strain mechanical

displacement relations, and the piezoelectric constitutive relations, respectively (Tiersten, H.F., 1963) (Slobodnik, A.J., Jr., 1976). Finally, the components of total time average electromechanical power flow are given by

$$\text{Time Average Power} = -\frac{1}{2} \int_0^\infty \text{Re} \left[T_{ij} \dot{u}_j^* \right] dx_3 + \frac{1}{2} \int_0^\infty \cdot \text{Re} \left[\varphi \dot{D}_i^* \right] dx_3 \qquad (10)$$

where the two terms are, respectively, the total complex mechanical and electrical power components.

2.2   Major Material Factors

The expressions derived above have been written into a computer program (Campbell, J.J., ... 1968) (Jones, W.R., ... 1969) which has been used extensively to study the SAW properties of many materials (Slobodnik, A.J., Jr., ... 1973a). Using that program it is possible to calculate three of the four material factors of interest in this paper: surface acoustic wave velocity, piezoelectric coupling constant, and electro-mechanical power flow angle.

Beside being necessary in the calculation of the temperature coefficient of time delay, as shown below, the SAW velocity influences the choice of substrate material depending upon the application. For delay lines and other devices requiring large time delays and small package size, low velocity materials are attractive. On the other hand, for high frequency devices, high velocity materials are useful for reducing fabrication resolution requirements (Hartmann, C.S., ... 1975).

The piezoelectric coupling constant $\Delta v/v$ is defined (Campbell, J.J., ... 1968) for piezoelectric materials as the percentage difference in velocity between free surfaces ($\omega h = \infty$) and surfaces coated with an infinitesimally thin perfect conductor ($\omega h = 0$). The validity of this definition as an estimate of the surface wave coupling to an interdigital transducer has been demonstrated several times both experimentally (Collins, J.H., ... 1968) (Schulz, M.B., ... 1972) and theoretically (Ingebrigsten, K.A., 1969) (Coquin, G.A., ... 1967). Note that for the purposes of this paper piezoelectric coupling is described directly in terms of $\Delta v/v$ and not the k^2 coupling parameter. Coupling information is of utmost importance in the design of broad-band, low insertion loss devices. It has been shown that for a given maximum allowable amount of insertion loss, the maximum attainable fractional bandwidth is proportional to the square root of the coupling constant (Hartmann, C.S., ... 1975). Hence, because the coupling of lithium niobate is about forty times as large as that of quartz (See Fig.1), devices on lithium niobate have about six times the fractional bandwidth as devices on quartz, for the same amount of insertion loss (Hays, R.M., ... 1976). This explains the popularity of lithium niobate for broad-band applications.

The power flow angle φ is defined in Fig. 4 as the angle between the time average electromechanical power flow vector and the direction of propagation (phase velocity vector). Unless φ identically equals zero (defined as a pure-mode axis), the condition of beam steering is said to occur. The slope of the power flow angle, that is $\partial\varphi/\partial\theta$, is a highly important quantity. Its magnitude determines the amount of beam steering resulting from a given unintentional misalignment from a pure-mode axis, and its magnitude and sign determine the extent of surface-wave diffraction. A later section will deal with these subjects in detail.

The fourth important material factor is temperature sensitivity as measured by the first order temperature coefficient of time delay. This quantity is given by

$$\frac{1}{\tau} \frac{\partial\tau}{\partial T} = \left(\frac{1}{v_s}\right)^{-1} \frac{\partial}{\partial T}\left(\frac{1}{v_s}\right) = \frac{1}{l} \frac{\partial l}{\partial T} - \frac{1}{v_s} \frac{\partial v_s}{\partial T} = \alpha - \frac{1}{v_s} \frac{\partial v_s}{\partial T} \qquad (11)$$

where $1/v_s \partial v_s/\partial T$ is the velocity temperature coefficient, $\tau = 1/v_s$ is the delay time, l is the distance between two material points, and α is the coefficient of thermal expansion. To calculate the temperature coefficient of time delay, the velocity temperature coefficient in equation (11) is approximated by

$$\frac{1}{v_s} \frac{\partial v_s}{\partial T} \approx \frac{1}{v_s(25^\circ C)} \left[\frac{v_s(35^\circ C) - v_s(15^\circ C)}{20^\circ C} \right] \qquad (12)$$

and the program described above is used to calculate SAW velocities at the various temperatures. Temperature sensitivity is important in many applications. For example, it has been shown that the principal limitation on the application of surface wave encoders and decoders to multiple-access secure communications systems is the degradation of the peak-to-sidelobe ratio of the autocorrelation function due to temperature differences (Carr, P.H., ... 1972). Also, in SAW bandpass filters the temperature stability of the center frequency is a direct function of the temperature coefficient of time delay of the substrate.

Since SAW materials are anisotropic, the four quantities discussed above are generally calculated for various crystalline orientations as continuous graphical functions of either direction of propagation in the plane of a plate (plates), as functions of the direction of the plate normal (boules), or for simultaneous rotation of both the plate normal and direction of propagation (cylinders). As an example, figure 5 shows plots of these four quantities as a function of the direction of propagation for the X cut of berlinite (Carr, P.H., ... 1976a). They were obtained by fixing the Euler angles λ and μ at 90.0 degrees and varying the angle θ.

3.   PROPAGATION LOSS

The optimum design of a surface acoustic wave device requires an adequate knowledge of the many sources of overall device insertion loss. The relationship of insertion loss and fractional bandwidth to the piezoelectric coupling constant was discussed in the previous section. Other sources of device insertion loss include bidirectionality loss, electrical mismatch loss, parasitic transducer conduction loss, matching network loss, apodization loss, spurious mode generation loss, substrate propagation losses, and beam steering and diffraction. Discussions of the first six of these subjects can be found in references (IEEE Proceedings, 1976), (Hartmann, C.S., ... 1975), (IEEE Trans. MTT, 1973), and (Wagers, R.S., 1976).

Substrate propagation losses are considered in this section, and beam steering and diffraction will be discussed in the next section.

Total propagation loss is a superposition of three different mechanisms (Slobodnik, A.J., Jr.,....1970): (1) Interaction with thermally excited elastic waves. (2) Scattering by crystalline defects and surface scratches. (3) Energy lost to air adjacent to the surface. The first mechanism is an inherent crystalline property, the magnitude of which can be predicted using viscosity theories (King, P.J., ... 1969). The second is, of course, highly undesirable and, fortunately, can be made negligible by proper crystal growth and polishing techniques (Slobodnik, A.J., Jr., ... 1970). The final mechanism is caused by the surface wave being phase matched to a longitudinal bulk wave in the air which results in a leaky-wave phenomenon. This so-called air loading can be eliminated by vacuum encapsulation or minimized by the use of a light gas (Slobodnik, A.J., Jr., 1972).

Propagation losses can be determined by directly probing the acoustic energy with a laser (Slobodnik, A.J., Jr., ... 1970). In this method, the surface wave deflects a small fraction of the incident light, which is detected with a photomultiplier tube and measured with a lock-in amplifier. The deflected light is directly proportional to the acoustic power of the surface wave.

Air loading can be determined by placing delay lines in a vacuum system and reducing the pressure below 1 torr while monitoring the change in insertion loss. Vacuum attenuation is, of course, the difference between the total propagation loss in air and the air loading component.

Frequency dependence of vacuum attenuation and air loading for three popular SAW substrates (Slobodnik, A.J., Jr., ... 1970), (Slobodnik, A.J., Jr., ... 1972), (Budreau, A.J., ... 1971), are illustrated in Figs. 6 and 7. Note the approximate f^2 dependence of the vacuum attenuation and the linear dependence of the air loading. This allows an empirical expression for propagation loss to be derived from the data.

$$\text{Propagation Loss (dB/}\mu s) = (VAC)F^2 + (AIR)F \tag{13}$$

where F is in GHz. The coefficients VAC and AIR are tabulated for popular substrates in reference (Slobodnik, A.J.,Jr.,1976). Equation (13) would be used, for example, when designing filters having particular bandpass characteristics.

4. BEAM STEERING AND DIFFRACTION

4.1 Parabolic Diffraction Theory

Diffraction of surface waves is a physical consequence of their propagation and can vary considerably depending upon the anisotropy of the substrate chosen. In fact, it is the slope of the power flow angle which determines the extent of both diffraction and beam steering (Szabo, T.L., . 1973). There is an inherent tradeoff between these two important sources of loss.

A useful theory for calculating diffraction fields when the velocity anisotropy near pure-mode axes can be approximated by a parabola has been developed by Cohen (Cohen, M.G., 1967). By using a small angle approximation, he showed that for certain cases, the higher orders of the expression for the velocity could be neglected past the second order. That is,

$$\frac{v(\theta)}{v_0} \approx 1 + \frac{\gamma}{2} (\theta - \theta_0)^2 \tag{14}$$

where $\gamma = \partial\varphi/\partial\theta$ and θ_0 is the angular orientation of the puremode axis. By comparing these approximations to an exact solution for electromagnetic diffraction in uniaxially anisotropic media, Cohen showed that the diffraction integral reduces to Fresnel's integral with the following change

$$\hat{Z}' = \hat{Z} \mid 1 + \gamma \mid . \tag{15}$$

Szabo and Slobodnik (Szabo, T.L., ... 1973) introduced the absolute magnitude signs to account for those materials having $\gamma < -1$ and the hatted terms to stand for wavelength scaled parameters ($\hat{Z} = Z/\Lambda$). In other words, diffraction is either accelerated or retarded depending on the value and sign of γ. Excellent agreement (Szabo,T.L., ... 1973) between this parabolic theory and experiment has been obtained whenever a good parabolic fit to the velocity is possible. In some cases, however, a more general theory is required.

4.2 Angular Spectrum of Waves Diffraction Theory

In order to solve the most general homogeneous anisotropic problem, Kharusi and Farnell (Kharusi, M.S., ... 1971) applied the angular spectrum-of-waves technique to surface-wave diffraction. Their theory is valid for both the near and far fields, and for any direction including off-axis orientations. Its only limitation is the requirement of accurate knowledge of velocity values for the surface of interest. In implementing their theory the following integration is performed numerically for each field point:

$$A(X,Z) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin K_1 L/2}{K_1} \cdot \exp j\{(K_1 X + K_3(K_1)|Z|)\} \, dK_1 . \tag{16}$$

Here K_3 and K_1 are the projections of the wave-vector K along the Z and X axes, respectively, or in general, along directions perpendicular to and parallel to the transducer. The effect of introducing a laser probe in the profile measurements can be accommodated (Szabo,T.L., ... 1973) by inserting

$$\frac{\sin K_1 P/2}{K_1 P/2} \tag{17}$$

(in which P is the probe diameter) into the preceding integral (Eq. 16).

The real power of the exact anisotropic theory can be illustrated by its ability to predict even the fine structure of a diffraction pattern on a highly nonparabolic velocity surface, including profile asymmetry due to beam steering. An example is shown in Fig. 8. The case studied (Szabo,T.L., ... 1973) concerns

surface waves launched close to the 111-axis of 211-cut gallium arsenide at a frequency of 280 MHz. Transducer widths were $\hat{L}$ = 51. This orientation was chosen (Szabo, T.L., ... 1973) because the velocity is non-parabolic and changes very rapidly with direction. The first column of Fig. 8 shows profiles for waves propagating exactly along the pure-mode 111-axis, a direction corresponding to φ = 0. Also note that the smoothing effect of the laser probe has not yet been included ($\hat{P}$ = 0). For the second column a misalignment of $0.6°$ from the 111-axis has been introduced, and the waves begin to take on the asymmetric behavior and beam steering of the experimental measurements (shown in the right-hand column) obtained using the laser probe (Slobodnik, A.J., Jr., ... 1970) technique. The third column introduces the same amount of angular misalignment as column two but, unlike the previous columns, includes the effect of a laser probe diameter of $\hat{P}$ = 5.3 wavelengths. The agreement between this column and the experimental curves is excellent.

4.3   Quantitative Choice of Theory

The versatility of the exact angular spectrum of waves theory has been demonstrated; however, this approach is far more computationally complicated and costly than the parabolic theory. It also requires precise velocity surfaces as input data. Given a certain material, then, the designer must have guidelines from which he can choose the simplest appropriate theory.

The closeness of a given velocity surface to a parabolic curve can be determined by fitting the surface to a parabola and noting any deviation. In particular, second-order fits were obtained (Szabo, T.L., ... 1973) for various materials by using a least squares fit with relative velocity values computed to seven significant places within a range of $\pm 5°$ of $(\theta - \theta_o)$ = 0. The maximum deviation of the fit from the velocity surface can be defined in terms of the quantity $|\delta_M|$. For comparative purposes, this deviation is expressed as a percentage of the actual velocity and, for convenience, is multiplied by a factor of 10^5, i.e.,

$$|\delta_M| = \left|\frac{v_{fit} - v}{v}\right| \times 10^5. \qquad (18)$$

A complete study of diffraction loss using the exact theory on many velocity surfaces not perfectly parabolic resulted in the following conclusion. Anisotropy may be conveniently grouped into two categories - parabolic $(0 < |\delta_M| \lesssim 2.0)$ and nonparabolic $(2.0 \lesssim |\delta_M| < \infty)$. Higher order terms discarded in the approximation of (Eq. 14) become significant (Szabo, T.L., ... 1973) for nonparabolic surfaces.

However for velocity surfaces having $|\delta_M| \lesssim 2.0$, the parabolic diffraction theory yields highly accurate results. Thus for all materials meeting this criterion, diffraction patterns are exactly equivalent in form and merely scaled in distance by the factor $|1 + \gamma|$ allowing universal diffraction loss curves to be calculated (Szabo, T.L., ... 1973). One such curve shown in Fig. 9 is a plot of diffraction loss versus the parameter $(\hat{z}/\hat{L}^2)|1 + \gamma|$. This curve allows the determination of loss for any combination of transducer width and separation for all parabolic anisotropic velocity surfaces. It was calculated by integrating the complex acoustic amplitude over the aperture of the receiving transducer for identical unapodized input and output transducers (Szabo, T.L., ... 1973).

In the Fresnel region the loss never exceeds 1.6 dB, which is the loss at the far-field length, $\hat{z} = \hat{z}_F$ (where the final peak in the beam profile has started its descent to a far-field pattern). The distance and transducer width at which a given loss will occur can always be given in far-field lengths. For example, the 3-dB loss point is

$$\hat{z}_{3\ dB} = 1.769\hat{z}_F \qquad (19)$$

where now

$$\hat{z}_F = \frac{\hat{L}^2}{|1 + \gamma|}. \qquad (20)$$

In the far field, the loss mechanism is the spreading of the beam with a slope of 10 dB/decade. The far-field loss can be approximated by

$$Loss\ (dB) = -10\ \log\frac{\hat{z}}{\hat{z}_F}. \qquad (21)$$

4.4   Minimal Diffraction Cuts

One extremely important implication of the parabolic diffraction theory is that since it reduces to the isotropic theory scaled by the factor $|1 + \gamma|$, no diffraction spreading occurs for ideal parabolic surfaces having $\gamma = -1$. Material orientations approaching this ideal have, in fact, been discovered (Slobodnik, A.J.,Jr., ... 1973b). A set of experimental SAW profiles for the 40.04 minimal diffraction cut (MDC) on bismuth germanium oxide are presented in Fig. 10. Experimentally, diffraction is suppressed by a factor of 100. These MDC orientations are allowing a new class of highly apodized acoustic surface-wave filters and long-time-delay devices to be realized.

4.5   The Beam Steering Diffraction Tradeoff

As mentioned at the outset of this section, there is an inherent tradeoff between beam steering and diffraction. In anisotropic materials, beam steering occurs whenever transducers are misaligned from a pure-mode axis θ_o, even though they may be perfectly aligned with each other. Beam steering is the pulling away of the acoustic beam from the transducer propagation axis by an additional angle, $\varphi = \gamma(\theta - \theta_o)$, as shown in Fig. 4. Let us discuss this tradeoff in more detail.

Diffraction is a fixed phenomenon for a given material, while beam steering can be controlled by precise X-ray alignment at the expense of increased device cost. Both, however, influence the choice of SAW substrate (Slobodnik, A.J.,Jr. ... 1974). An example of how the combined loss of beam steering and diffraction varies among materials is illustrated (Slobodnik, A.J., Jr., ... 1974) in Fig. 11 where the loss is given as a function of γ.

For Fig. 11 the acoustic aperture is $\hat{L}$ = 80 wavelengths, the distance between input and output transducer is $\hat{Z}$ = 5000 wavelengths, and the misalignment from the desired pure-mode axis, or the beam steering (BS) angle, is BS $\hat{\chi}$ = 0.1°. In order to use these data for practical situations, it is only necessary to insert the slope of the power flow angle appropriate to the type and cut under consideration. It is also useful to note that $\hat{Z}$ = tf; where t is the time delay and f the frequency of the device of interest.

Several important features can be noted with reference to Fig. 11. Diffraction loss goes to 0 for those materials having γ = - 1.0 and, as expected, the combined loss curve agrees exactly with the beam steering loss curve. Those materials having γ = 0 correspond to locally isotropic cases and beam steering goes to 0. Here, diffraction accounts for the total loss. Diffraction loss alone is symmetric about γ = - 1.0 and beam steering loss about γ = 0, while the combined curve is clearly nonsymmetric. Universal beam steering plus diffraction loss curves are not possible.

The results illustrated in Fig. 11 are of major importance in choosing a material for a particular application. For example, where diffraction is potentially a very serious problem, as in highly apodized filters, a material having γ = -1.0 would be most desirable.

Figure 12 illustrates (Slobodnik, A.J., Jr., ... 1974) combined beam steering and diffraction loss versus the time-delay-frequency parameter $\hat{Z}$. It is of interest to point out that the loss is very high for the $\hat{Z}$ = 75 000 curve near γ = - 1.0. For this large distance beam steering is very important, especially for narrow undiffracted beams, and some beam spreading is to be desired. (The same is true if inaccurate X-ray orientation must be tolerated.) Since $\hat{Z}$ is proportional to frequency (for fixed time delay), Fig. 12 also illustrates why beam steering and diffraction are considered UHF and microwave frequency design problems. Significant losses and material tradeoff considerations exist at the higher frequencies and, of course, also for very long time delays.

5.   HIGH COUPLING, TEMPERATURE COMPENSATED MATERIALS

5.1   Introduction

A good deal of attention has recently been devoted to the development of SAW devices having broad bandwidth, low insertion loss, and temperature independent performance characteristics. As stated in the Introduction, neither quartz nor lithium niobate is adequate for such devices. This has resulted in a search for substrate materials that are temperature compensated and have piezoelectric coupling greater than that of ST cut quartz (Schulz, M.B., ... 1970). One material which has been extensively studies is lithium tantalate (LiTaO$_3$). Although this material is temperature compensated for volume waves (Detaint, J., ... 1976), no temperature compensated cut is known to exist for surface acoustic waves. Nevertheless, as Fig. 1 shows, the piezoelectric coupling and temperature coefficient of time delay for lithium tantalate represent a reasonable compromise between the high coupling and poor temperature coefficient of lithium niobate and the very low coupling of temperature compensated quartz. Of particular interest is the minimum diffraction cut (MDC) of lithium tantalate, which has 1/20th the diffraction spreading of an isotropic material (Slobodnik, A.J., Jr., ... 1975). The LiTaO$_3$ minimum diffraction cut has very low coupling to bulk waves, a very important property for low spurious response filters (Slobodnik, A.J., Jr., ... 1975) and delay lines (Carr, P.H., ... 1976b).

A major contribution to the search for improved materials has been the development of a phenomenological model (Newnham, R.E., 1973) which explains why known materials are temperature compensated. According to that model, temperature compensated materials possess either of the following anomalous properties: (1) a positive temperature coefficient of velocity or elastic constant or (2) a negative coefficient of thermal expansion. Quartz, for example, is temperature compensated because the temperature coefficient of C_{66}, the elastic constant for shear propagation along the Z axis, is positive (Newnham, R. E., 1973).

5.2   Berlinite

Berlinite, (AlPO$_4$), is structurally similar to quartz with larger piezoelectric constants. A recent investigation (Chang, Z.P., ... 1976) has shown that for bulk waves berlinite is indeed temperature compensated along orientations similar to the AT and BT cuts of quartz, but with 2.5 times larger piezoelectric coupling. Motivated by these results for bulk waves and the fact that the temperature coefficient of one of berlinite's elastic constants is positive (Chang, Z.P., ... 1976). Studies have recently been conducted to investigate the behaviour of surface acoustic waves on berlinite (Carr, P.H., ... 1976a), (O'Connell, R.M., ... 1977a), (O'Connell, R.M., ... 1977b). That investigation has produced several temperature compensated cuts with more than four times the piezoelectric coupling of ST cut quartz (Carr, P.H., ... 1976a) (O'Connell, R.M., ... 1977a). High coupling, temperature compensated orientations have been found for doubly rotated cuts as well as for singly rotated cuts. In order to investigate the SAW behaviour of berlinite, the theoretical computer model described in Section 2 was used in conjunction with the data of reference (Chang, Z.P., ... 1976) to calculate the SAW velocity, the piezoelectric coupling, the electromechanical power flow angle, and the first order temperature coefficient of time delay for several standard crystallographic cuts (O'Connell, R.M., ... 1977b).

Results for the X cut of berlinite are shown in Figure 5. These and other initial results (Carr, P.H., ... 1976a) showed that berlinite is temperature compensated with four times the piezoelectric coupling of ST cut quartz. However, all of those initially reported cuts had non-zero electromechanical power flow angles, and consequently suffer from beam steering. Subsequent calculations produced two singly rotated cuts and two doubly rotated cuts, all of which have zero electromechanical power flow angles (O'Connell, R.M., ... 1977a). The most promising of the singly rotated cuts is the X axis boule 80.4° cut, a direct analog of the ST cut of quartz. The two cuts are compared in Table 1. Note that the piezoelectric coupling, $\Delta V/V$, of the X-axis boule 80.4° cut is more than four times as large as that of ST quartz, a distinct advantage. Other calculations have been made to investigate the behaviour of pseudo surface acoustic waves on berlinite (Jhunjhunwala, A., ... 1976a).

Table 1 also shows that the slope of the power flow angle, $\partial\phi/\partial\theta$, is larger for the X axis boule 80.4° cut of berlinite than it is for the ST cut of quartz. According to the theory of SAW diffraction discussed in

Section 4, this means that ST cut quartz has better diffraction properties than the X axis boule 80.4°
cut of berlinite. The desire to find temperature compensated cuts of berlinite having better diffraction
properties than the singly rotated 80.4° cut motivated consideration of doubly rotated cuts. In particular,
the μ = 90.0 plane of an orthogonal coordinate system having the three Euler angles λ, μ, and θ as its basis
was carefully searched. The μ = 90.0 plane was of particular interest because it contains four of the stan-
dard crystallographic cuts, including the X cut and Z axis cylinder, for which temperature compensated ori-
entations were found earlier (Carr, P.H., ... 1976a).

The results are shown in Figure 13. The dashed and solid curves represent, respectively, the loci of Euler
angles for which the electromechanical power flow angle and the temperature coefficient of time delay are
zero. As can be seen in the blown up portion of the Figure, the loci intersect in a total of twelve places
throughout the plane. Because of crystal symmetry, however, only two of the points are independent, and
those circled in Figure 13 are listed in Table 1. Notice that while they have about the same piezoelectric
coupling as the singly rotated 80.4° cut, the slopes of their power flow angles are smaller than those of
either the 80.4° cut or the ST cut of quartz, giving them the added advantage of less diffraction spreading.

5.3 β-Eucryptite

Another material which has been the subject of recent attention (Barsch, G.R., ... 1975) is β-eucryptite,
(Li Al SiO$_4$). This material is interesting because of its large negative coefficient of thermal expansion
in the direction of the hexagonal C axis. According to the above-mentioned phenomenological model for ex-
plaining why certain materials are temperature compensated, this may give rise to temperature compensation
even though the temperature coefficients of the elastic constants are all negative (Barsch, G.R., ... 1975).
Using the theoretical computer model and data from references (Barsch, G.R., ... 1975), (Hummel, F.A.,1951),
(Schulz, H., 1974), and (Bohm, H., 1975), calculations of the SAW properties of β-eucryptite produced both
singly rotated and doubly rotated temperature compensated cuts. A singly rotated cut was found (O'Connell,
R.M., ... 1977c) at X cut 69°, and is listed in Table 1 which shows that although the piezoelectric coupling
for this cut is almost twice as large as that of ST quartz, it has the disadvantage of an 18 degree elec-
tromechanical power flow angle.

As was done in the case of berlinite, doubly rotated cuts were considered also, and a temperature compen-
sated cut having a zero electromechanical power flow angle was found in the λ = 0.0° plane (O'Connell, R.M.,
... 1977c), as shown in Figure 14. As can be seen in the Figure, the loci intersect in a total of four
places throughout the plane. Again, because of crystal symmetry, only one of the points is independent,
and it is listed in Table 1 where it can be seen that, unfortunately, the piezoelectric coupling of this
doubly rotated cut is only about half as large as that of ST quartz. Perhaps the most attractive feature
of this material is that it has the highest SAW velocity of all the temperature compensated materials list-
ed in Table 1, 3662 m/sec.

5.4 Other Temperature Compensated Materials

The sulfosalts are a class of materials of the form Tl$_3$BX$_4$, where B can be V, Nb, or Ta, and X can be S or
Se. Recent calculations have shown that at least two of these materials are temperature compensated with
significantly larger piezoelectric coupling than berlinite (Weinert, R.W., ... 1975), (Isaacs, T.J., ...
1976). One particular cut of Tl$_3$ VS$_4$, for example, has four times the piezoelectric coupling of berlinite
(Weinert, R.W., ... 1975). As shown in Table 1, however, this cut has the disadvantage of a rather large
electromechanical power flow angle, about -17 degrees. Another cut of the same material and one of
Tl$_3$ Ta Se$_4$, having zero electromechanical power flow angles, have also been found (Jhunjhunwala, A., ...
1976b). As the data in Table 1 shows, the piezoelectric coupling of these cuts is not as large as that of
the first cut discussed, but it is still more than twice as large as that of berlinite. The table also
shows that the SAW velocities of the sulfosalts are only about 1/3 as large as that of berlinite. This is
a disadvantage for high frequency applications, but an advantage for long delay lines and low frequency SAW
filters.

A composite material, consisting of a film of silicon dioxide on lithium tantalate, has also been shown to
be temperature compensated (Parker, T.E., ... 1975). This material has, as shown in Table 1, a very small
electromechanical power flow angle, a piezoelectric coupling of about .007, and a relatively large SAW
velocity. The most attractive feature of the material is that its second order temperature coefficient of
time delay is nearly an order of magnitude smaller than that of ST cut quartz. Despite these positive
attributes, the composite has several drawbacks due to the SiO$_2$ film, including: (1) its thickness must be
very accurately controlled, (2) it is very lossy at high frequencies, and (3) it is dispersive.

Besides those materials which have been shown to be temperature compensated, there are several others which,
for various reasons, may prove to be. For example, nepheline, (KAlSiO$_4$) (NaAlSiO$_4$)$_3$, should have tempera-
ture compensated orientations because the temperature coefficients of the elastic constants C_{11} and C_{66}
are positive (Bonczar, L.J., ... 1975). Also, lead potassium niobate (Pb$_2$KNb$_5$O$_{15}$), which has piezoelectric
coupling factors of the order 0.7 for bulk waves, shows promise of being temperature compensated because it
possesses opposite signs for the temperature coefficients of the bulk wave resonance frequency for different
vibrational modes (Barsch, G.R., ... 1976).

6. FAST VELOCITY MATERIALS

Fast velocity materials, such as beryllium oxide and aluminum nitride, are of great interest for extending
the upper frequency limit of SAW devices. Also, the high thermal conductivity of these materials results
in a high power handling capacity; for example, continuous wave power levels of 1W produced no damage to a
transducer operating at 1 GHz (Budreau, A.J., ... 1974a). Beryllium oxide single crystals have been grown
in centimeter sizes (Hagon, P.J., ... 1971). The velocity of surface acoustic waves is 6500 m/sec. The
very weak piezoelectric coupling (1/10th that of quartz) and the strong coupling to volume waves make this
material unattractive compared to aluminum nitride thin-films-on-sapphire, which have unusually low cou-
pling to volume waves and a piezoelectric coupling up to six times that of ST-quartz. The velocity of the
films varies from about 6000 m/sec for very thin films to 5500 m/sec for 0.7 wavelength thick films (Liu,
J.K., ... 1975a). Also their temperature coefficients vary from 47 ppm for a very thin film to 25 ppm for

a film 0.57 wavelengths thick (Liu, J.K., ... 1975b). Single crystals of aluminum nitride have been grown in sizes up to only several millimeters (Dugger, C.0., 1974), (Slack, G.A., ... 1976); larger crystals are needed for measuring the unknown elastic constants.

Velocity nonuniformities from one side to the other of a 1 cm wide sample have been observed to be of the order of 0.5% for thin films of AlN grown on sapphire by chemical vapor deposition (Budreau, A.J., ...1974b). This velocity difference is of course most critical for narrowband SAW filter applications. Another problem has been the lack of reproducibility of the piezoelectric coupling from sample to sample. Liu, et. al. (Liu, J.K., ... 1975a), (Liu, J.K., ... 1975b) have shown that this can be due to misalignment of the pyramidal grains of the as-grown thin films. The granularity is so severe that the films must first be polished to avoid scattering the surface acoustic waves. Initial results with films grown by reactive rf sputtering show that they do not have this granularity; thus, this method of growing films warrants further investigation (Shuskus, A.J., ... 1974).

7.   CONCLUSIONS

Clearly, the task of choosing a substrate material for a SAW device is not trivial. There are many factors which must be considered, most of which have been discussed here. Fortunately, the relative importance of these factors varies with the specific device and appplication at hand. It is hoped that the topics discussed herein will provide the design engineer with a reasonable understanding of the many factors involved in SAW device substrate choice, and with a useful set of guidelines for weighing the relative importance of those factors in a specific design.

REFERENCES

BARSCH, G. R. AND SPEAR, K.E., 1975, "Temperature Compensated Piezoelectric Materials," AFCRL Report No. TR-75-0609, Contract No. F19628-75-C-0085.

BARSCH, G.R. AND SPEAR, K.E., 1976, "Temperature Compensated Piezoelectric Materials," RADC Report No. TR-76-184, Contract No. F19628-75-C-0085.

BOHM, H., 1975, "Dielectric Properties of β-Eucryptite," Phys.Stat.Sol., Vol. (a) 30, p 531.

BONCZAR, L.J., AND BARSCH, G.R., 1975, "Elastic and Thermoelastic Constants of Nepheline," J.Appl.Phys., vol 46, pp 4339-4340.

BUDREAU, A.J. AND CARR, P.H., 1971, "Temperature Dependence of the Attenuation of Microwave Frequency Elastic Surface Waves in Quartz," J.Appl.Phys., vol. 18, pp. 239-241.

BUDREAU, A.J., SILVA, J.H., AND LAKER, K.R., 1974a "Microwave Frequency SAW Filters on Aluminum Nitride-on-Sapphire and Beryllium Oxide," 1974 Ultrasonics Symp. Proc. IEEE Cat No 74896-1SU, p 299.

BUDREAU, A.J., LAKER, K.R., AND CARR, P.H., 1974b, "Compact Microwave Acoustic Surface Wave Filter Bank for Frequency Synthesis," Proceedings of the Symposium on Optical and Acoustical Micro-Electronics, Polytechnic Institute of New York, p 471.

BUDREAU, A.J., KEARNS, W.J., AND CARR, P.H., 1975, "State-of-the-art in Microfabrication of SAW Devices," 1975 Ultrasonics Symposium Proceedings, IEEE Cat. No. 75 CHO 994-4SU, pp 458-462.

CAMPBELL, J.J. AND JONES, W.R., 1968, "A Method for Estimating Optimal Crystal Cuts and Propagation Directions for Excitation of Piezoelectric Surface Waves," IEEE Trans. Sonics Ultrason., vol. SU-15, pp. 209-218.

CARR, P.H., DeVITO, P.A., AND SZABO, T.L., 1972, "The Effect of Temperature and Doppler Shift on the Performance of Elastic Surface Wave Encoders and Decoders," IEEE Trans.Sonics Ultrason., vol.SU-19, pp. 357-367.

CARR, P.H. AND O'CONNELL, R.M., 1976a, "New Temperature Compensated Materials with High Piezoelectric Coupling," Proc. of the 30th Annual Symposium on Frequency Control, pp 129-131.

CARR, P.H., FENSTERMACHER, T.E., SILVA, J.H., KEARNS, W.J., AND STIGLITZ, M.R., 1976b, "SAW Delay Line With All Spurious 70 dB Down," 1976 Ultrasonics Symposium Proceedings, IEEE Cat.No. 76 CH1120-5SU, p 459.

CHANG, Z.P. AND BARSCH, G.R., 1976, "Elastic Constants and Thermal Expansion of Berlinite," IEEE Trans.on Sonics and Ultrasonics, Vol. SU-23, pp 127-135.

COHEN, M.G., 1967, "Optical Study of Ultrasonic Diffraction and Focussing in Anisotropic Media," J.Appl. Phys., vol. 38, pp 3821-3828.

COLLINS, J.H., GERARD, H.M., AND SHAW, H.J., 1968, "High-Performance Lithium Niobate Acoustic Surface Wave Transducers and Delay Lines," Appl.Phys.Lett., vol. 13, pp. 312-313.

COQUIN, G.A. AND TIERSTEN, H.F., 1967, "Analysis of the Excitation and Detection of Piezoelectric Surface Waves in Quartz by Means of Surface Electrodes," J.Acoust.Soc.Amer., vol. 41, pp. 921-939.

DETAINT, J. AND LANCON, R., 1976, "Temperature Characteristics of High Frequency Lithium Tantalate Plates," Proc. of the 30th Annual Symposium on Frequency Control, p 132.

DUGGER, C. O., 1974, "The Synthesis of Aluminum Nitride Single Crystals," Mat.Res.Bull., vol 9; p 331.

GOLDSTEIN, H., 1950, Classical Mechanics, Reading, MA: Addison-Wesley.

HAGON, P.J. AND AUSTERMAN, S.B., 1971, "Acoustic Waves in Beryllium Oxide Crystal," Appl.Phys.Letters, Vol 18, p 102.

HARTMANN, C.S., BELL, D.T., JR., AND ROSENFELD, R.C., 1975, "Impulse Model Design of Acoustic Surface-Wave Filters," IEEE Trans. Microwave Theory Tech., Vol. MTT-21, pp 162-175.

HAYS, R.M., AND HARTMANN, C.S., 1976, "Surface-Acoustic-Wave Devices for Communications," Proc. IEEE, Vol 64, pp 652-671.

HUMMEL, F.A., 1951, "Thermal Expansion Properties of Some Synthetic Lithia Minerals," J.Am.Ceram.Soc., Vol 34, p 235.

IEEE Proceedings, 1976, "Special Issue on Surface Acoustic Waves," Vol 64, No. 5.

IEEE Trans. Microwave Theory Tech., 1973, Special Issue on Microwave Acoustic Signal Processing, Vol MTT-21, No. 4.

INGEBRIGTSEN, K.A., 1969, "Surface Waves In Piezoelectrics," J.Appl.Phys., vol. 40, pp 2681-2686.

ISAACS, T.J., AND WEINERT, R.W., 1976, "Crystal Growth and Properties of Tl_3BX_4 Crystals for Acoustic Surface-Wave and Bulk Acoustic Devices," Journal of Electronics Materials, Vol 5, pp 13-22.

JHUNJHUNWALA, A., VETELINO, J.F., AND FIELD, J.C., 1976a, "Berlinite, A Temperature Compensated Material for Surface Acoustic Wave Applications," Proc. of the 1976 Ultrasonics Symposium, pp. 523-527.

JHUNJHUNWALA, A., VETELINO, J.F., AND FIELD, J.C., 1976b, "Temperature Compensated Cuts with Zero Power Flow in Tl_3VS_4 and Tl_3 Ta Se_4," Electronics Letters, Vol. 12, pp 683-684.

JONES, W.R., CAMPBELL, J.J., AND VEILLEUX, S.L., 1969, "Theoretical Analysis of Acoustic Surface Waves," Hughes Aircraft Co., Fullerton, CA, Final Report F19628-69-0132, unpublished.

KHARUSI, M.S. AND FARNELL, G.W., 1971, "Diffraction and Beam Steering for Surface-Wave Comb Structures on Anisotropic Substrates," IEEE Trans. Sonics Ultrason., vol. SU-18, pp. 35-42.

KING, P.J. AND SHEARD, F.W., 1969, "Viscosity Tensor Approach to the Damping of Rayleigh Waves," J.Appl. Phys., vol. 40, pp. 5189-5190.

LIU, J.K., LAKIN, K.M., AND WANG, K.L., 1975a, "Growth Morphology and SAW Measurements of AlN Films on Sapphire," J.Appl.Phys., vol 46, p 3703.

LIU, J.K., STOKES, R.B., AND LAKIN, K.M., 1975b, "Evaluation of AlN Films on Sapphire for SAW Applications," 1975 Ultrasonics Symposium Proceedings, IEEE Cat. No. CHO 994-4SU, p 234.

NEWNHAM R.E., 1973, "Elastic Properties of Oxides and the Search for Temperature Compensated Materials," AFCRL Report No. TR-73-0220, Contract No. F19628-73-C-0108.

O'CONNELL, R.M., AND CARR, P.H., 1977a, "High Piezoelectric Coupling, Temperature Compensated Cuts of Berlinite, Al PO_4, for SAW Applications," IEEE Trans. on Sonics and Ultrasonics.

O'CONNELL, R.M., AND CARR, P.H., 1977b, "New Materials for Surface Acoustic Wave (SAW) Devices," Optical Engineering.

O'CONNELL, R.M. AND CARR, P.H., 1977c, "Temperature Compensated Cuts of Berlinite and β-eucryptite for SAW Devices," 31st Annaul Frequency Control Proceedings (to be published).

PARKER, T.E. AND SHULZ, M.B., 1975, "Stability of SAW Controlled Oscillators," Proc. of the 1975 Ultrasonics Symposium, pp 261-263.

RAYLEIGH, LORD, 1885, "On Waves Propagated Along the Plane Surface of an Elastic Solid," Proc.London Math. Soc., vol 17, pp. 4-11.

SCHULZ, H., 1974, "Thermal Expansion of Beta Eucryptite," J.Am.Ceram.Soc., Vol 57, p 313.

SCHULZ, M.B., MATSINGER, B.J., AND HOLLAND, M.G., 1970, "Temperature Dependence of Surface Acoustic Wave Velocity on ∝ Quartz," J.Appl.Phys., Vol 41, pp 2755-2765.

SCHULZ, M.B. AND MATSINGER, J.H., 1972, "Rayleigh-wave Electromechanical Coupling Constants," Appl.Phys. Lett., vol. 20, pp 367-369.

SHUSKUS, A.J., REEDER, T.M., AND PARADIS, E.L., 1974, "RF-Sputtered Aluminum Nitride Films on Sapphire," Appl.Phys.Letters, vol 24, p 155.

SLACK, G.A. AND McNELLY, T.R., 1976, "Growth of High Purity AlN Crystals," J.of Crystal Growth, Vol 34, p 263.

SLOBODNIK, A.J., JR., CARR, P.H., AND BUDREAU, A.J., 1970, "Microwave Frequency Acoustic Surface-Wave Loss Mechanisms on $LiNbO_3$," J.Appl.Phys., vol 41, pp 4380-4387.

SLOBODNIK, A. J.,JR., 1972, "Attenuation of Microwave Acoustic Surface Waves Due to Gas Loading," J.Appl. Phys., vol. 43, pp 2565-2568.

SLOBODNIK, A.J., JR., AND BUDREAU, A.J., 1972, "Acoustic Surface Wave Loss Mechanisms on $Bi_{12}GeO_{20}$ at

Microwave Frequencies," J.Appl.Phys., vol. 43, pp 3278-3283.

SLOBODNIK, A.J., JR., CONWAY E.D., AND DELMONICO, R.T., 1973a, "Microwave Acoustics Handbook, Vol. 1A, Surface Wave Velocities," AFCRL, Hanscom AFB, MA 01731, TR-73-0597, unpublished.

SLOBODNIK, A.J., JR., SZABO, T.L., 1973b, "Minimal Diffraction Cuts for Acoustic Surface Wave Propagation on $Bi_{12}GeO_{20}$," J.Appl.Phys., vol. 44, pp 2937-2941.

SLOBODNIK, A.J., JR., SZABO, T.L., 1974, "Design of Optimum Acoustic Surface Wave Delay Lines at Microwave Frequencies," IEEE Trans. Microwave Theory Tech., vol. MMT-22, pp 458-462.

SLOBODNIK, A.J.,JR., FENSTERMACHER, T.E., KEARNS, W.J., ROBERTS, G.A., AND SILVA, J.H., 1975, "A Minimal Diffraction Lithium Tantalate Substrate for Contiguous SAW Butterworth Filters," 1975 Ultrasonics Symposium Proceedings, IEEE Cat. No. 75 CHO 994-4SU, p 405.

SLOBODNIK, A.J.,JR., 1976, "Surface Acoustic Waves and SAW Materials," Proc. IEEE, Vol 64, pp 581-595.

SZABO, T.L. AND SLOBODNIK, A.J., JR., 1973, "The Effect of Diffraction on the Design of Acoustic Surface Wave Devices," IEEE Trans. Sonics Ultrason., vol. SU-20, pp 240-251.

TIERSTEN, H.F., 1963, "Thickness Vibrations of Piezoelectric Plates," J.Acoust.Soc.Amer., vol.35, pp 53-58.

WEINERT, R.W. AND ISAACS, T.J., 1975, "New Piezoelectric Materials Which Exhibit Temperature Stability for Surface Waves," Proc. of the 29th Annual Symposium on Frequency Control, pp 139-142.

WHITE, R.M. AND VOLTMER, F.W., 1965, "Direct Piezoelectric Coupling to Surface Elastic Waves," Appl.Phys. Lett., vol. 7, pp. 314-316.

WAGERS, R.S., 1976, "Spurious Acoustic Responses in SAW Devices," Proc. IEEE, Vol. 64, pp 699-702.

TABLE 1

Temperature Compensated Cuts of Various Materials

| MATERIAL	ORIENTATION	EULER ANGLES			POWER FLOW ANGLE ϕ (DEG)	SLOPE OF POWER FLOW ANGLE $(\partial\phi/\partial\theta)$	$\Delta v/v_\infty$ (X 10^{-2})	SAW VELOCITY (m/sec)
		λ	μ	θ				
QUARTZ (SiO_2)	ST CUT	0	132.75	0	0.0	0.378	0.058	3158
BERLINITE ($AlPO_4$)	X AXIS BOULE 80.4°	0	80.4	0	0.0	0.901	0.245	2751
	DOUBLY ROTATED, A	76.8	90	11.5	0.0	0.372	0.250	2756
	DOUBLY ROTATED, B	79.7	90	15.5	0.0	0.221	0.247	2758
β-EUCRYPTITE (β-LiAlSiO$_4$)	X CUT 69°	90	90	69	18	--	0.100	3662
	DOUBLY ROTATED	0	57	62	0.0	0.32	0.035	3258
Tl_3VS_4	(110) CUT 70°	-45	90	70	-17	--	1.0	900
	(110) CYLINDER 24°	45	24	90	0.0	--	0.617	1010
Tl_3TaSe_4	(110) CYLINDER 54°	45	54	90	0.0	--	0.508	879
$SiO_2/LiTaO_3$	Y CUT, Z PROP	0	90	90	0.0	--	0.7	3455
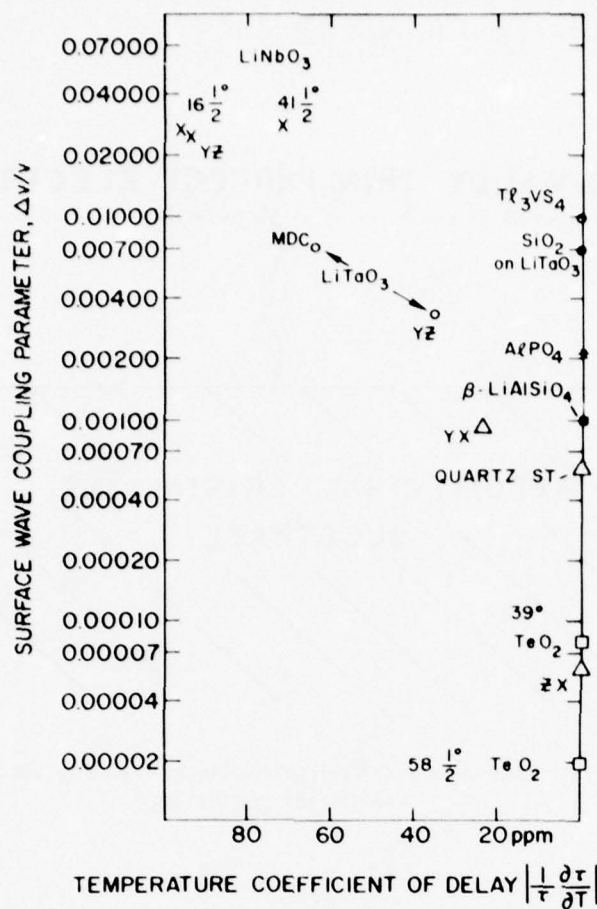
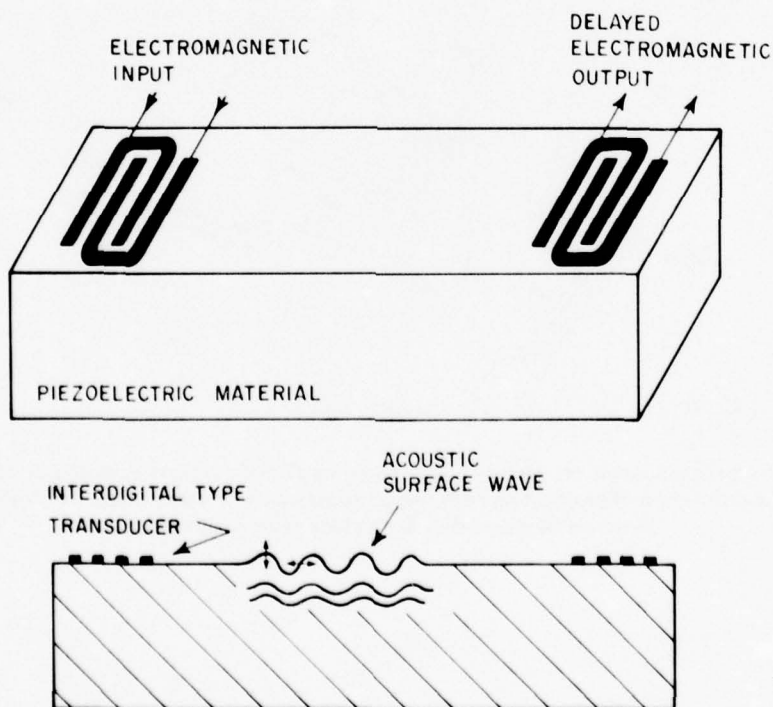Fig.1   Temperature coefficient of time delay versus piezoelectric coupling for various SAW materials.

Fig.2   Schematic representation of the generation and propagation of a surface acoustic wave.

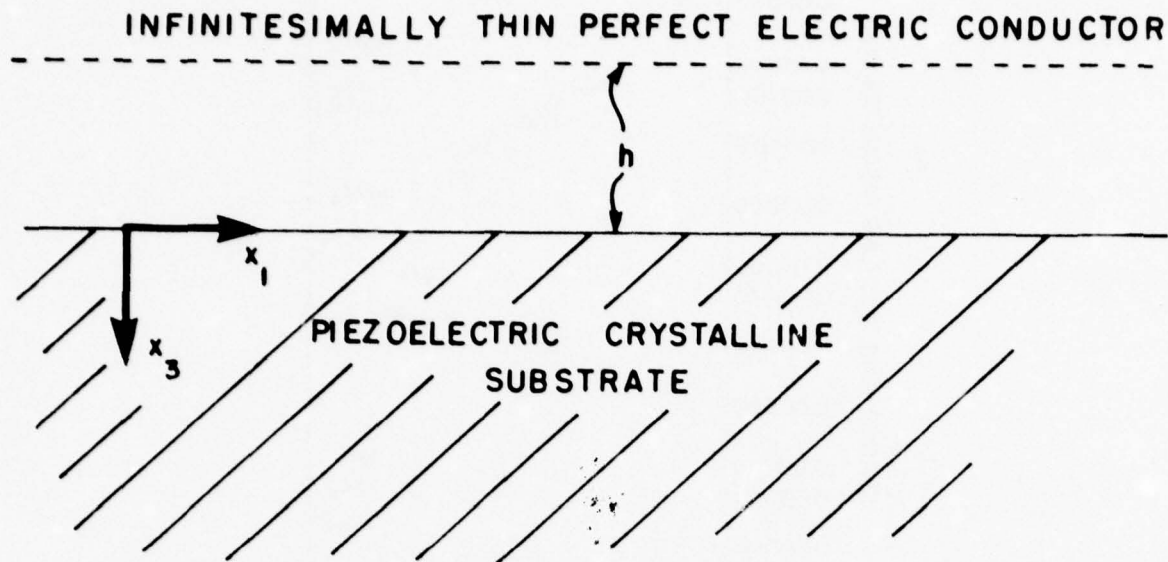# INFINITESIMALLY THIN PERFECT ELECTRIC CONDUCTOR

Fig.3 Illustration of the coordinate system used to define SAW propagation. The shorting plane will be necessary when computing the quantity $\Delta v/v$.
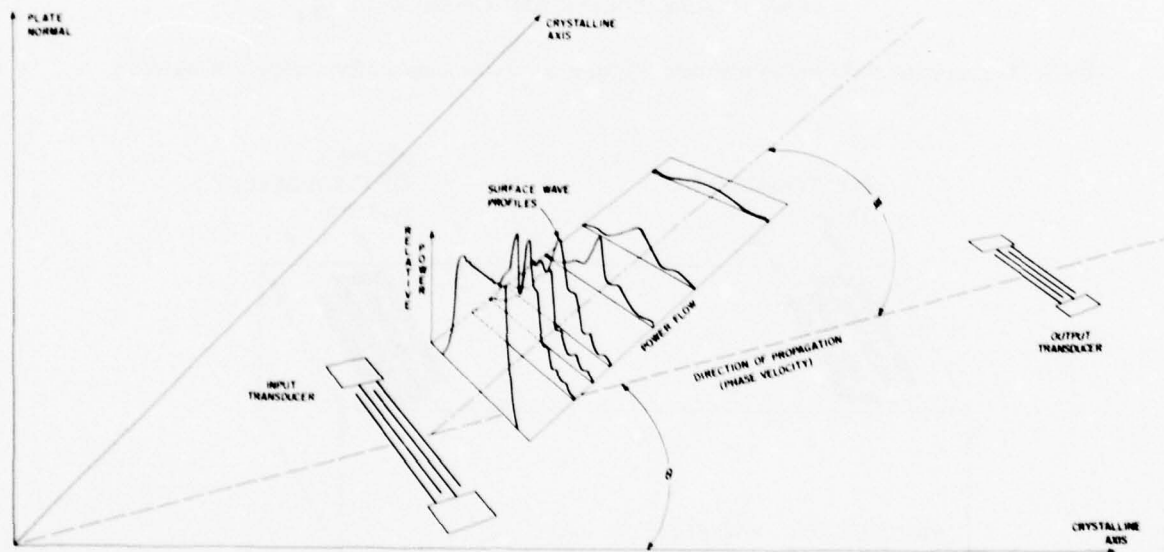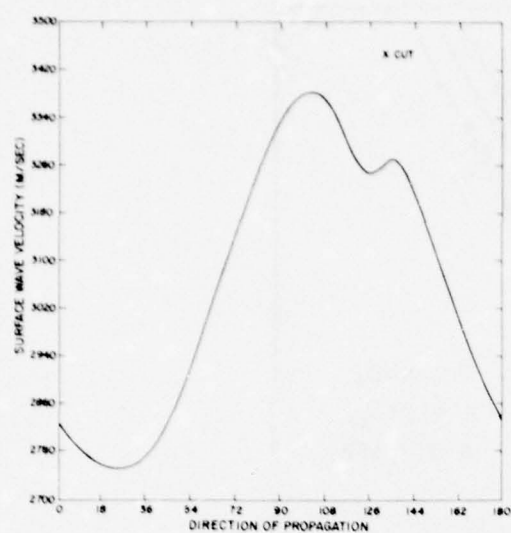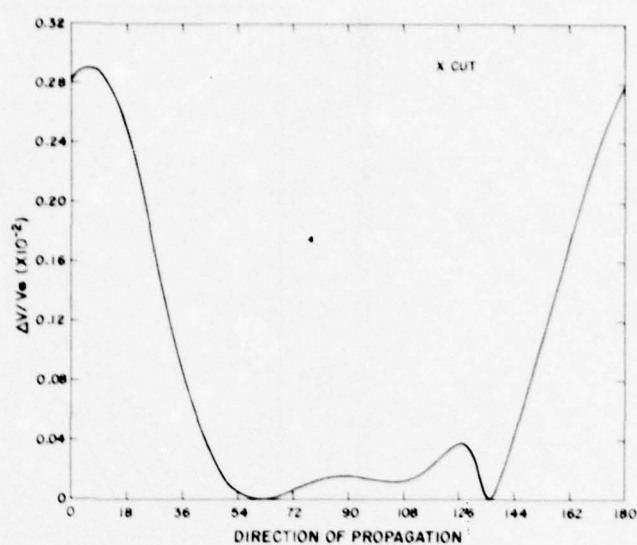

Fig.4 Schematic representation of the profiles of a propagating acoustic surface wave on a crystalline substrate. Angle θ defines direction of propagation with respect to reference crystalline axis, and angle φ defines deviation of power flow from phase velocity direction.
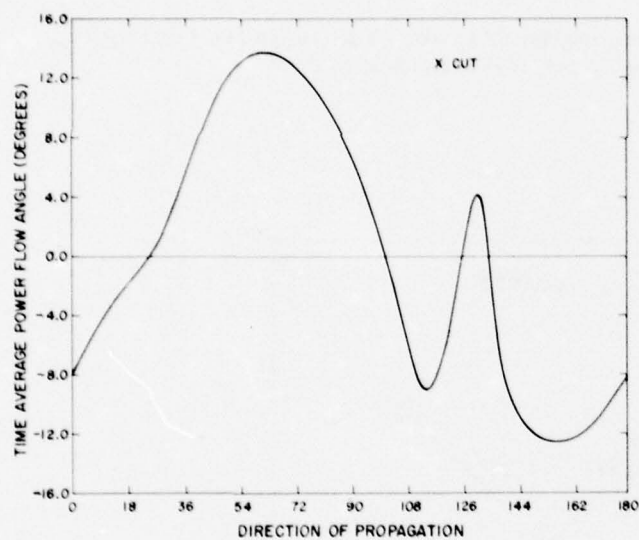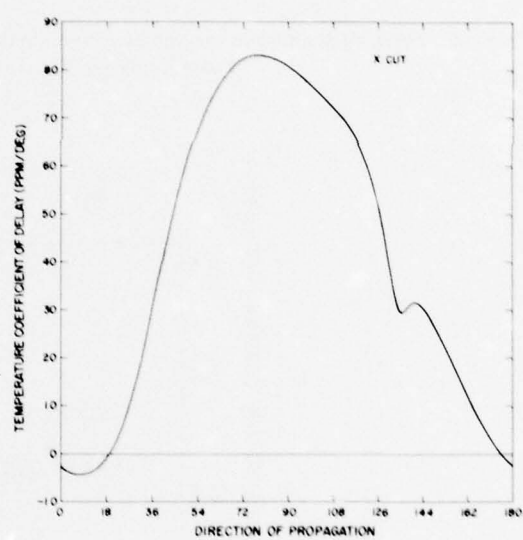
Fig.5 The variation of (a) SAW velocity, (b) piezoelectric coupling, (c) electromechanical power flow angle, and (d) temperature coefficient of time delay for X-cut berlinite.
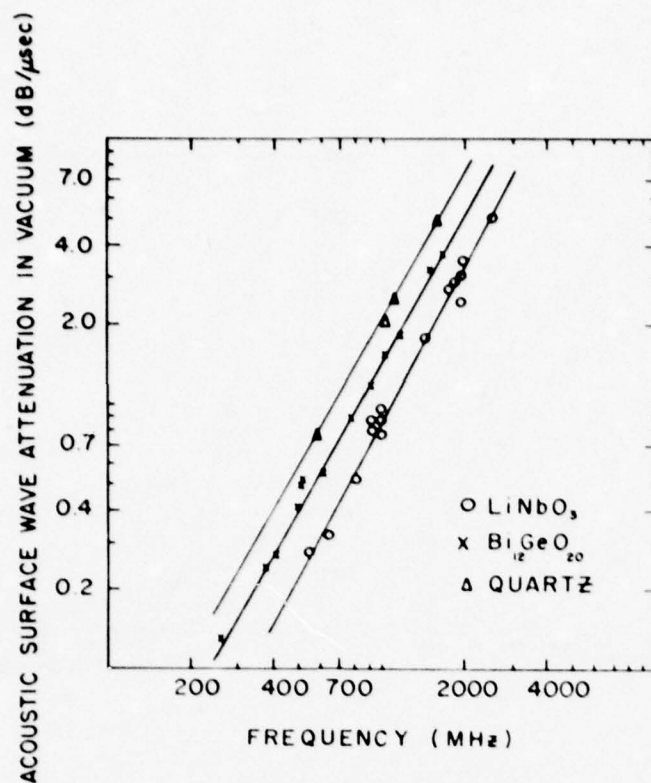
Fig.6 SAW attenuation in vacuum as a function of frequency for YZ LiNb0$_3$, 001, 110 and 111, 110, Bi$_{12}$Ge0$_{20}$, and YX quartz. Experimental slopes are all approximately f^2 .
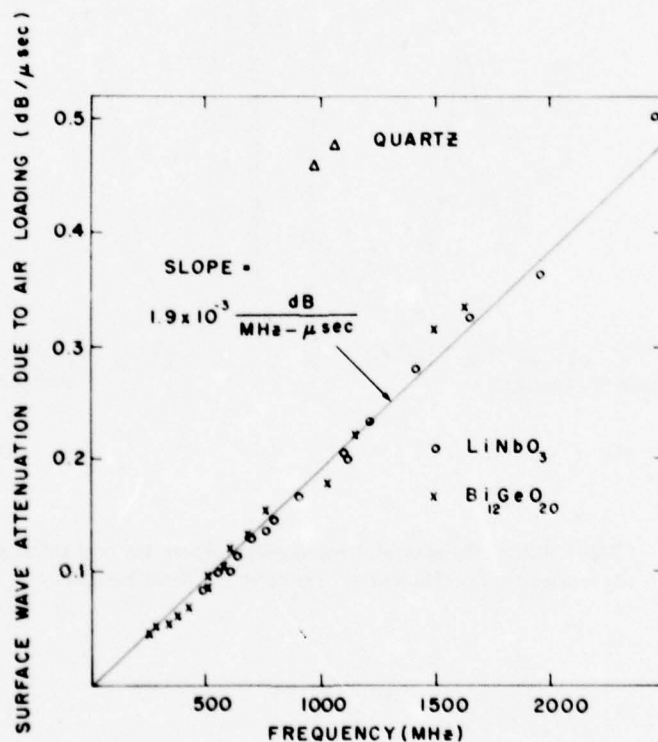
Fig.7 SAW attenuation due to air loading as a function of frequency for materials listed in Fig.6. It is interesting to note nearly identical results for LiNb0$_3$ and Bi$_{12}$Ge0$_{20}$.
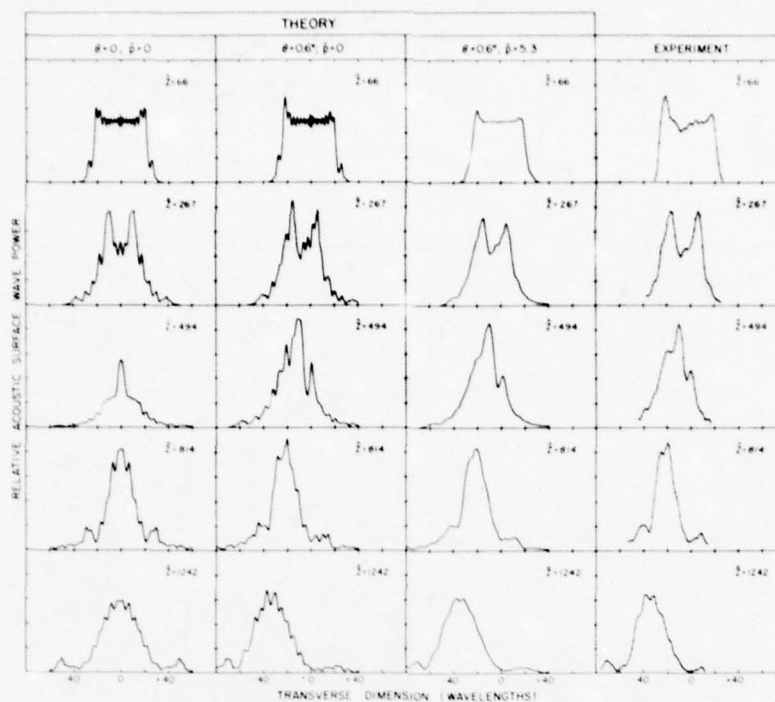
Fig.8 Theoretical and experimental surface-wave profiles illustrating diffraction near 111-axis of 211-cut gallium arsenide. Ż indicates distance for propagation in wavelengths from input transducer, θ gives the misorientation from 111-axis, and Ṗ is laser probe diameter in acoustic wavelengths.
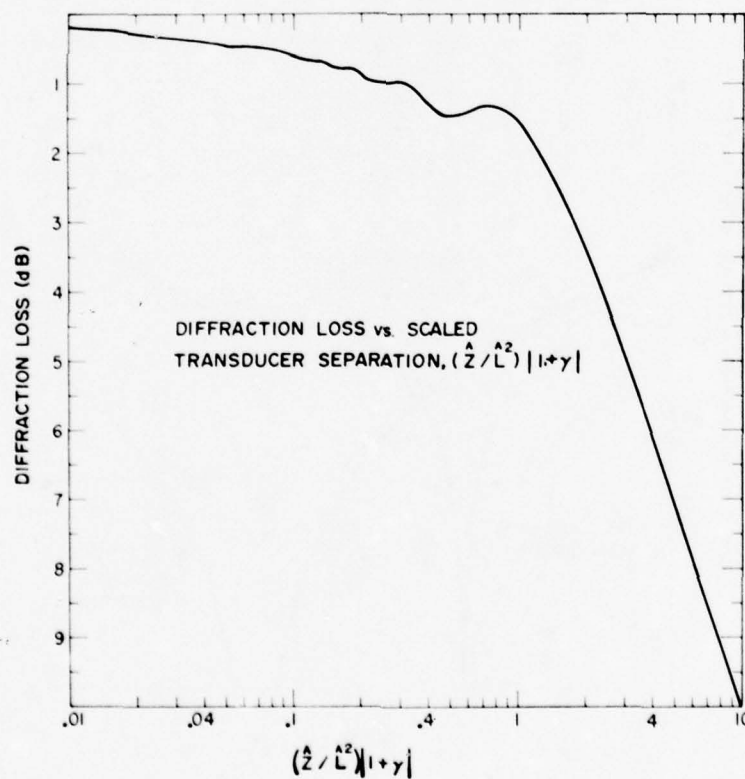
Fig.9 Universal diffraction loss curve for all parabolic materials as a function of $(\hat{Z}/\hat{L}^2)|1 + \gamma|$. To convert to the actual distance in wavelengths on horizontal scale simply insert L̇ , width of your transducer in wavelengths, and γ appropriate to your material.
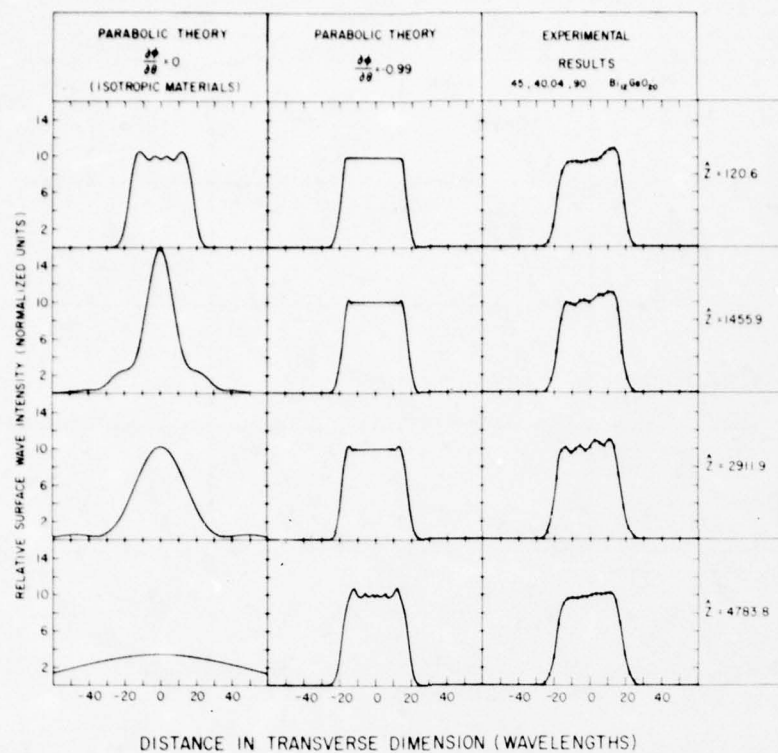
Fig.10   Illustration of the two orders of magnitude diffraction suppression achieved using the 40.04 $Bi_{12}GeO_{20}$ minimal diffraction cut. An acoustic aperture of $\hat{L} = 40.56$ wavelengths was used.
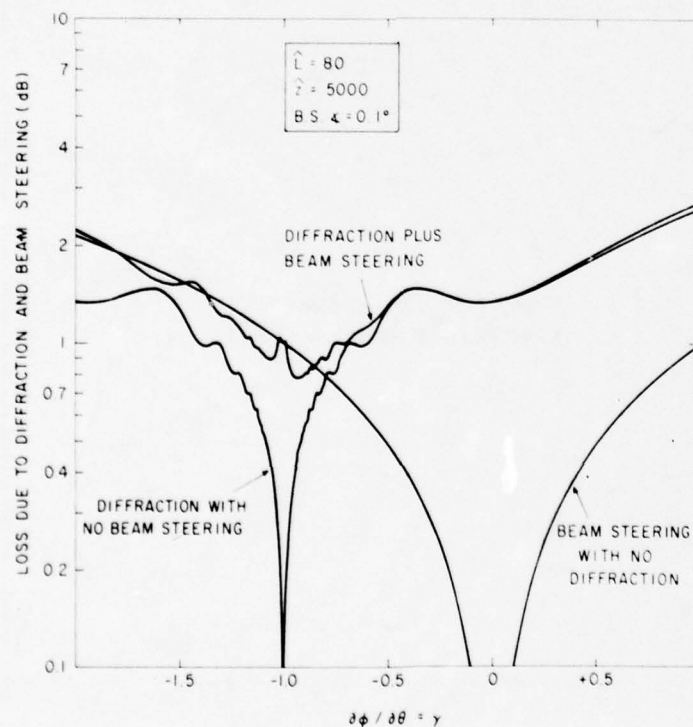
Fig.11   Loss due to diffraction and beam steering as a function of slope of power flow angle for parabolic materials. $\hat{L}$ represents width of transducer in wavelengths, $\hat{Z}$ the distance between transducers in wavelengths, and BS $\ll$ the beam steering angle (defined as misalignment of center line between transducers from desired pure-mode axis).
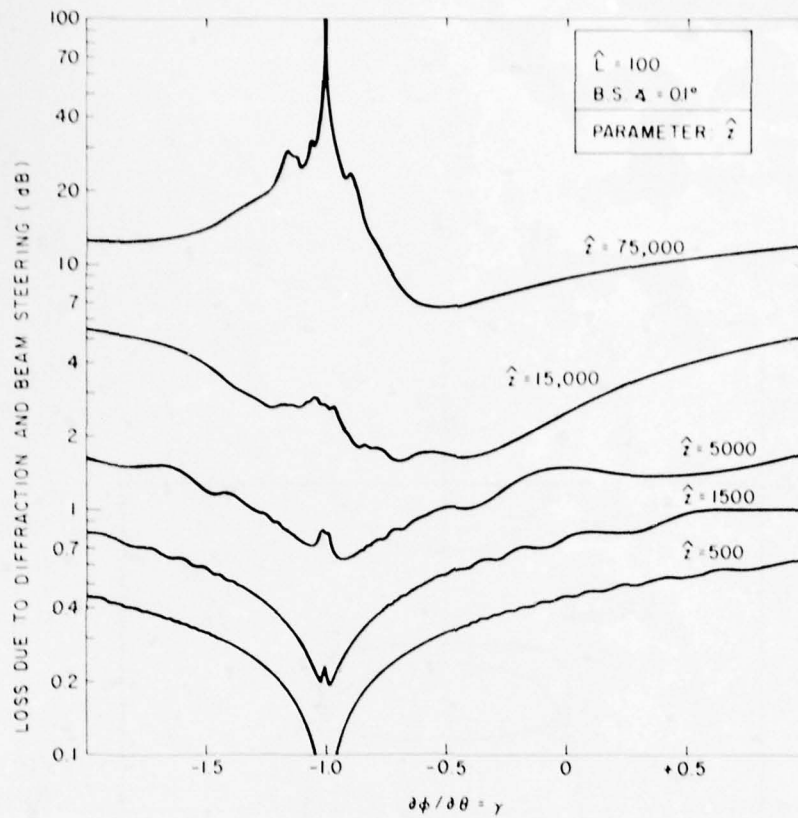
Fig.12   Loss due to diffraction and beam steering as a function of slope of power flow angle with distance in wavelengths between transducers, $\hat{Z}$ as parameter.
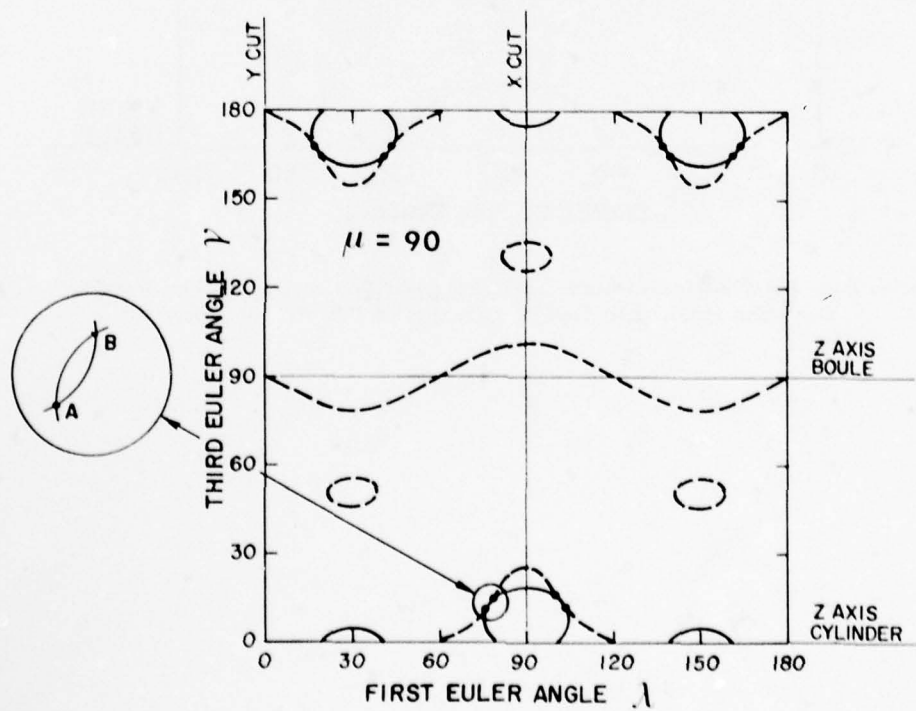
Fig.13   Loci of Euler angles having zero electromechanical power flow angle (dashed lines) and zero temperature coefficient of time delay (solid lines) in the $\mu = 90.0$ plane of berlinite.
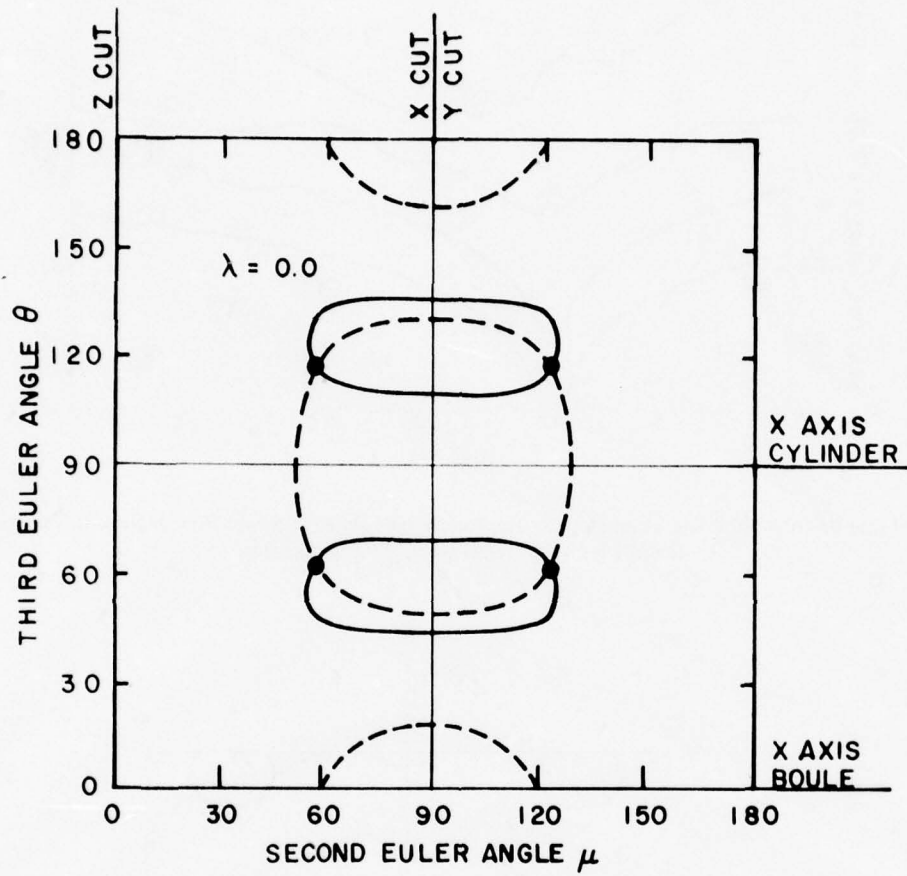
Fig.14 Loci of Euler angles having zero electromechanical power flow angle (dashed lines) and zero temperature coefficient of time delay (solid lines) in the $\lambda = 0.0$ plane of β-eucryptite.

DISCUSSION

P.Tournois

As far as T1$_3$ VS4 is concerned do you have infrared transparency and permittivity figures?

Author's Reply

They have been measured by Westinghouse, but I do not have the figures.

Voles

Can you comment on the frequency dispersion at all? IBM have found that in practice there is a very slight dispersion when there is a damaged layer.

Comments – E.Stern

In grating devices, energy storage in grooves introduces a dispersion as large as $1 : 10^4$. This dispersion is anticipated in our design equations and it typically causes minimal problems in structures up to 10^5 wavelength long.

E.Stern

Could you recommend material for 100 °C temperature variations?

Author's Reply

The temperature variation of a SiO_2 thin film on $LiTaO_3$ is significantly less than that of ST quartz over a 100 °C temperature range.

MICROWAVE SURFACE-ACOUSTIC-WAVE COMPONENTS

P. HARTEMANN

THOMSON-CSF Research Center

Domaine de Corbeville - 91401 Orsay (France)

SUMMARY

Various surface-acoustic-wave components operating at frequencies greater than 1 GHz have been realized using an electron beam pattern generator. Fabrication process are described in this paper and the main features of devices are reported. Simple and tapped delay lines operating at about 1.3 GHz with a 3 dB bandwith close to 500 MHz were performed with lithium niobate substrates. The maximum delay time is equal to 10 µs .Large bandwidth coded lines and dispersive delay lines were also tested. Narrow bandwidth filters were realized using sampled interdigital transducers deposited on surfaces of lithium niobate and quartz substrates. A relative bandwidth of 8×10^{-4} has been obtained. Oscillators have been implemented with such filters up to 2 GHz, the substrates being ST cut quartz plates.

1.   INTRODUCTION

Surface-acoustic-wave (SAW) devices are now currently used for signal processing at frequencies lower than about 500 MHz. But in order to increase the bandwidth of transducers and the fundamental frequency of oscillators ,frequencies greater than 1 GHz must be reached. The transducers of SAW devices are generally made of fingers with a quarter-wavelength width that is about 0.9 µm at 1 GHz. Sophisticated pattern generator must be employed for obtaining such a submicron-size finger width. An other limitation of interdigital transducers is due to the great electrical resistance of a narrow finger, the thickness of fingers being only about 600 Å for the highest frequency and the ohmic loss of a large bandwidth transducer composed of a few number of fingers is large. Moreover propagation loss limits also the upper frequency of SAW devices. At 1 GHz this loss for lithium niobate and quartz is close to 1 dB/µs and 3 dB/µs respectively and it is about proportionnal to the squared frequency (SLOBODNIK A.J., 1973). Then fabrication process used for obtaining very narrow fingers by electron-beam lithography are described in this paper and features of some SAW components operating at frequencies greater than 1 GHz are reported.

2.   FABRICATION   PROCESS

The electron beam pattern generator fabricated by THOMSON-CSF is employed for obtaining very narrow fingers (CAHEN O. et al., 1972). The substrate is fixed on a (X,Y) table, the position of which is measured using two Michelson interferometers with an accuracy equal to 0.04 µm. The desired position of the impact spot of the electron beam is reached with motors moving the table and coils deflecting the beam. The global accuracy of this masking machine is about 0.1 µm over a scan field of 25 cm². The spot of the electron beam can sweep a 500 µm by 500 µm square without moving the table. This machine is completely digitalized, the elementary spot displacement being 0.125 µm.

The piezoelectric substrates are exposed directly to the electron beam and a static charge is accumulated because they are insulators. This charge must be eliminated in order to avoid electron beam distorsion; this is done by a metallic coating. Thus the 200 Å thick aluminium film intended for suppressing the static charges may be deposited on the free surface of the substrate or on the 4000 Å thick P.M.M.A. resist coating. In the first case the parts of this film located between fingers are ion etched at the end of the fabrication process. (see the figure 1). In the second case the electrons reach the resist through the aluminium film which is chemically etched before the development. (see the figure 2). Except for these last precautions, the conventional lift-off technique is employed for both cases in order to deposit aluminium fingers with a thickness close to 800 Å. The second process is simpler than the first because ion etching is not necessary and short-circuits between fingers are less frequent. The smallest width so obtained for the transducer finger has been 0.3 µm as shown in the figure 3. In this case the centre frequency of the delay line has been 2.7 GHz. The insertion loss has been excessive (80 dB) mainly because the ohmic loss of the transducer has been large   , the resistance of a finger being about 300 Ω.

3.   CHARACTERISTICS OF MICROWAVE SAW COMPONENTS
3.1   Delay lines

Y-cut, Z-propagating lithium niobate substrates have been used for realizing simple or tapped delay lines. Parts of frequency responses corresponding to lines with a delay of 2.2µs and 10µs respectively are shown in the figure 4. The width of transducer fingers is 0.58 µm and the length 250 µm. The main features of these two delay lines are presented in the table I.

	$T_{µs}$	N	P_{dB}	$F_{o\ GHz}$	$B_{1\ MHz}$	$B_{2\ MHz}$
Line no. 1	2.2	4	43	1.34	540	720
Line no. 2	10	6	62	1.25	280	390

Table I

T : delay - N : number of fingers for each transducer
P : insertion loss with matching networks - F_0 : centre frequency
B_1 : 3 dB bandwidth - B_2 : 6 dB bandwidth

The centre frequency of the line n° 2 are lower than that of the line n° 1 because of the propagation loss variation versus frequency.

Transducers with a variable finger spacing (dispersive transducer) have been performed for reducing insertion loss. Using a dispersive transducer made of 37 fingers a loss of 31.5 dB for a 1.1 μs delay has been obtained with a 3 dB bandwidth of 500 MHz, the other transducer being a 4 finger standard transducer. In this case the delay variation is 10 ns for a 700 MHz range and the finger width changes from 0.47 μm to 0.76 μm along the dispersive transducer.

A 10 taps delay line has been fabricated with an incremental delay of 1 μs. The launching transducer and the 10 identical receiving transducers are made of 10 and 4 fingers respectively. The frequency response of each output is shown in the figure 5. The insertion losses are between 26.6 and 52.6 dB according to the number of the output. The discrepencies occuring in the shapes of frequency responses are due to the propagation loss variation within the bandwidth. The average centre frequency and bandwidth are about 1.4 GHz and 400 MHz respectively.

3.2 Coded lines

A line coded according to a 13 chip Barker code was tested, the centre frequency being 1.5 GHz. The chip duration is 6.5 ns. The coded transducer is made of 4 finger elementary transducers. The finger number of the conventional transducer is 20. The 3 dB width of the compressed pulse is about 6 ns.

3.3 Dispersive delay lines

Dispersive delay lines with a bandwidth close to 500 MHz have been performed using dispersive transducers at 1.3 GHz (WEGLEIN R.D. et al., 1973) and reflective arrays of grooves at 1 GHz (WILLIAMSON R.C. et al., 1973). In the case of dispersive transducers, the compression ratio is limited because the number of narrow fingers becomes excessive for a dispersive delay greater than about 500 ns and short-circuits between fingers are frequent. This number of short-circuits is reduced using dispersive transducers operating at an harmonic frequency. Transducers with split fingers operate well at the third harmonic. (BRISTOL T.W. et al., 1972). At 1.5 GHz the finger width is 0.9 μm for such a transducer instead of 0.6 μm for a conventional transducer. Thus a line made of two identical dispersive transducers with split fingers was tested at 1.3 GHz. The delay variation is 212 ns for a 300 MHz range according to a linear law (BT=63.6). The number of split fingers is equal to 95 for each transducer, the finger length being 250 μm. The insertion loss is close to 30 dB with impedance matching networks, the highest frequencies being the most delayed.

3.4 Filters

Narrow bandwidth filters have been realized with a transducer pattern previously tested at lower frequencies (HARTEMANN P., 1971)   . Sampled transducers have been used in order to avoid a great number of fingers. They are made of elementary transducers separated by a spacing equal to an integer of the wavelength corresponding to the centre frequency Fo. This number is equal to N_1 and N_2 for the launching and reveiving transducer respectively as shown in the figure 6. For instance this couple of integers is equal to 45 and 55. In this case the frequency response exhibits main peaks separated by a frequency interval equal to $F_0/5$ that is 300 MHz for a centre frequency Fo of 1.5 GHz. Such a filter was performed. The two transducers consist of 21 and 17 four finger elementary transducers. The fingers are 0.58 μm wide. The spacing between two adjacent elementary transducers is 105 μm and the total transducer length 2.1 mm for one transducer ; 120 μm and 2.05 mm for the other. In order to obtain a transducer with such length, the combs were realized in four parts, the table of the masking machine being moved between each exposure. The quality of the joints between the differents parts is very good. The frequency response of this filter using a (Y, Z) cut lithium niobate as the substrate shows two peaks at 1.237 GHz and 1.540 GHz in this frequency range (see the figure 7). The bandwidth is about 1.1 MHz and the insertion loss without impedance matching about 25 dB. The shape of these peaks is in good agreement with the theoretical predictions.

The same transducers have been deposited on a ST cut quartz substrate. In this case the peak frequencies are 1.12 GHz, 1.40 GHz and 1.70 GHz, the 3 dB bandwidth being about 1 MHz. The loss is 28 dB with impedance matching. The very great accuracy of the electron beam pattern generator at THOMSON-CSF has been essential for obtaining a successful operation of this kind of filter.

3.5 Oscillators

It is well known that a delay line looped with an amplifier compensating the line loss may oscillate if the phase shift along the loop is equal to   integer times 360° (LEWIS M.F., 1973). The narrow bandwidth filters described in the previous paragraph have been used in order to realize such oscillators. ST cut quartz is the more useful substrate because its temperature stability is very good. In this case the oscillating frequency may correspond to either of the transfer function peaks, according to the frequency of the impedance matching. The spectra of these transmitted waves at 1.12 GHz, 1.40 GHz and 1.70 GHz are shown in the figure 8 for a 2 GHz range. These spectra are very clean with no detectable spurious lines.

A 2 GHz oscillator has been performed using the same kind of transducer. The insertion loss of the delay line has been about 48 dB with a ST cut quartz substrate. The transducers are composed of 7 elementary transducers with 8 fingers. Using a greater number of elementary transducers the loss can be reduced.

At 1.01 GHz using 15 and 13 elementary transducers made of 8 fingers the insertion loss has been 20.5dB with a ST cut quartz substrate and the spectrum obtained by multiplying this frequency up to 11 GHz is shown in the figure 9 .The peak level is close to 50 dB above noise for a 300 Hz analysis bandwidth.

4. CONCLUSION

Microwave acoustic surface wave components have been fabricated by exploiting the accuracy of an electron beam pattern generator. The narrowest finger width that we have obtained is equal to 0.3 μm and the corresponding frequency is close to 2.7 GHz. However a centre frequency of 1.5 GHz is more practical, the finger being 0.6 μm wide. In this frequency range the SAW components are reliable and their features are attractive for applications. So large bandwidth simple and tapped delay lines with a maximum delay of 10 μs have been achieved. Great data rate coded lines, large bandwidth dispersive delay lines and narrow bandwidth filters have been also obtained. Moreover delay line oscillators operating at a fundamental frequency larger than 1 GHz have been performed. Their use presents a great advantage :the frequency multiplier required for reaching the 10 GHz range is very simplified. In the future SAW resonators operating at about 1 GHz will be performed. They will be used for implementing filters with low insertion and oscillators with a very good spectral purity.

REFERENCES

BRISTOL T.W.,JONES W.R., SNOW P.B and SMITH W.R.1972 "Applications of double electrodes in acoustic surface wave device design "Ultrasonics Symposium Proceedings, 343

CAHEN O., SIGELLE R., TROTEL J., 1972 " Automatic control of an electron beam pattern generator" Electron and Ion beam science and technology, 5th International Conference Proceedings, 92

HARTEMANN P., 1971 "Narrow bandwidth Rayleigh Wave filters" Electronics Letters 7, 674

LEWIS M.F., 1973 "Some aspects of SAW oscillators" Ultrasonics Symposium Proceedings, 344

SLOBODNIK A.J., 1973 "A review of material tradeoffs in the design of acoustic surface wave devices at VHF and microwave frequencies " IEEE Transactions in Sonics and Ultrasonics SU-20, 315

WEGLEIN R.D., WAUK M.T. and NUDD G.R., 1973 " 500 MHz bandwidth surface wave pulse compression filter" Ultrasonics Symposium Proceedings, 482

WILLIAMSON R.C., DOLAT V.S., and SMITH H.I., 1973 " L-band reflective - array compressor with a compression ratio of 5120", Ultrasonics Symposium Proceedings, 490
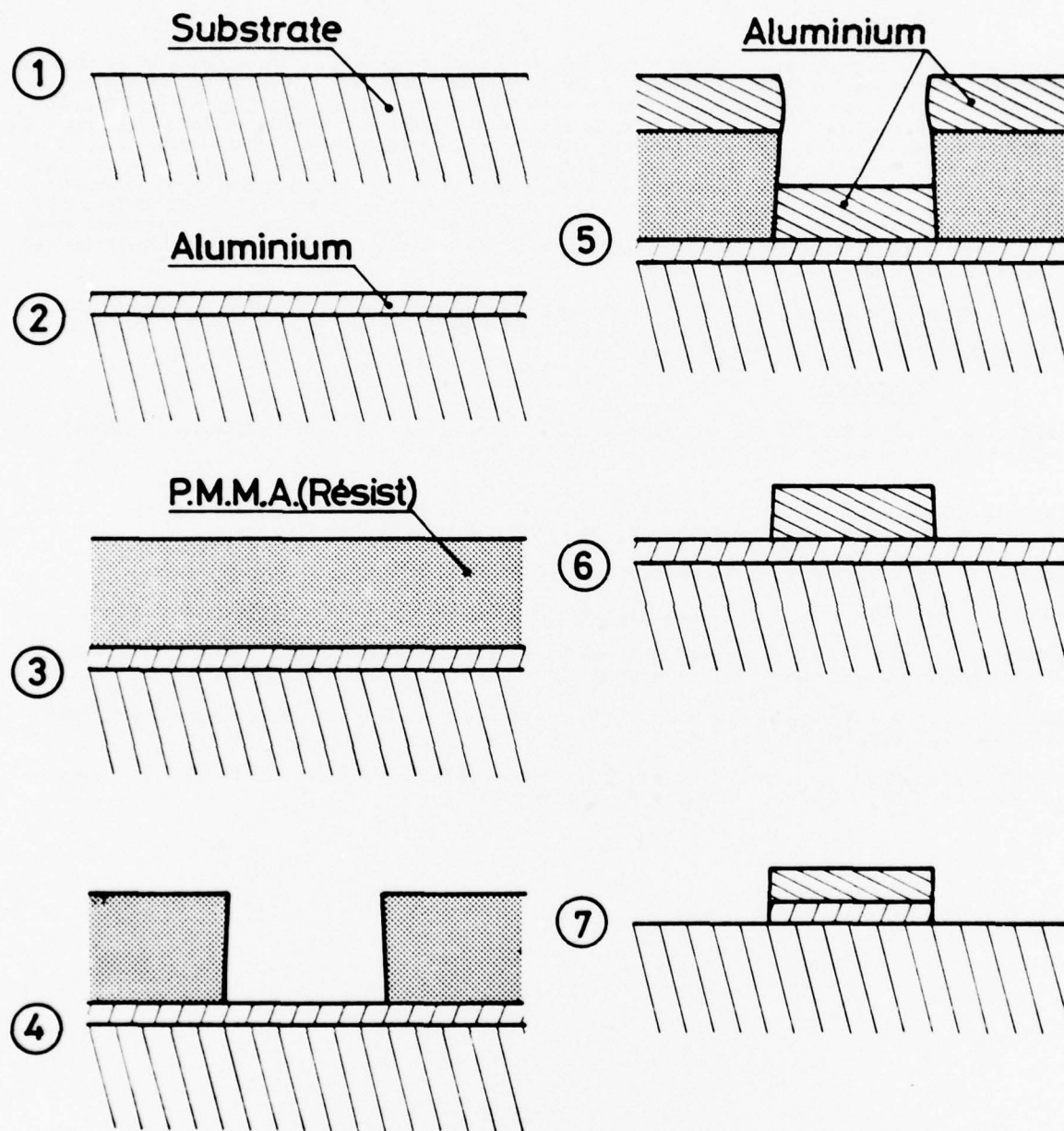
fig. 1   first fabrication process used for obtaining very narrow fingers with an electron beam pattern generator. The first aluminium film is ion etched between the states 6 and 7. The coating of resist (poly (Methyl Methacrylate)) is 4000 Å thick.
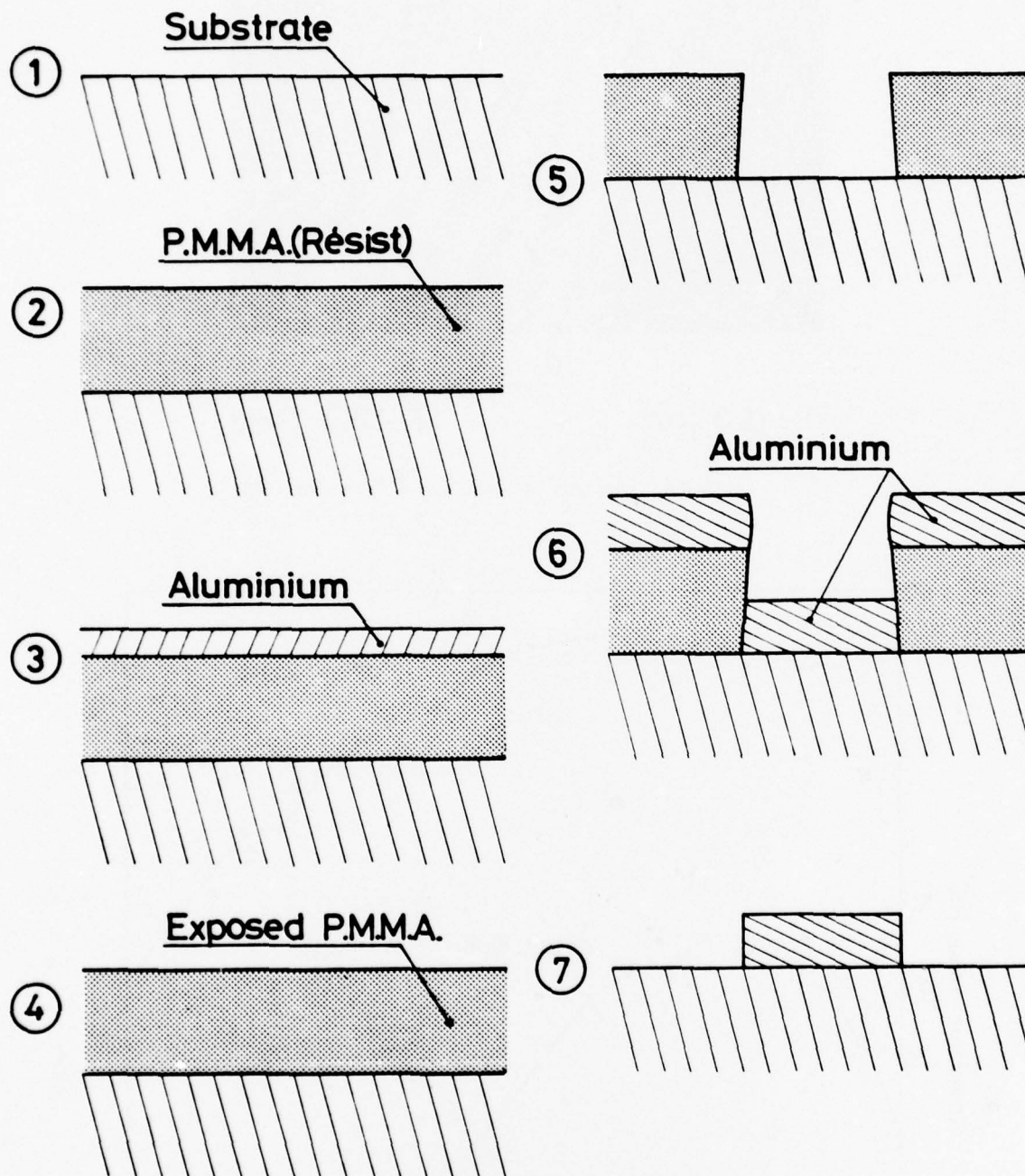
fig. 2   second fabrication process used for obtaining very narrow fingers with an electron beam pattern generator. The first aluminium film is chemically etched between the states 3 and 4.
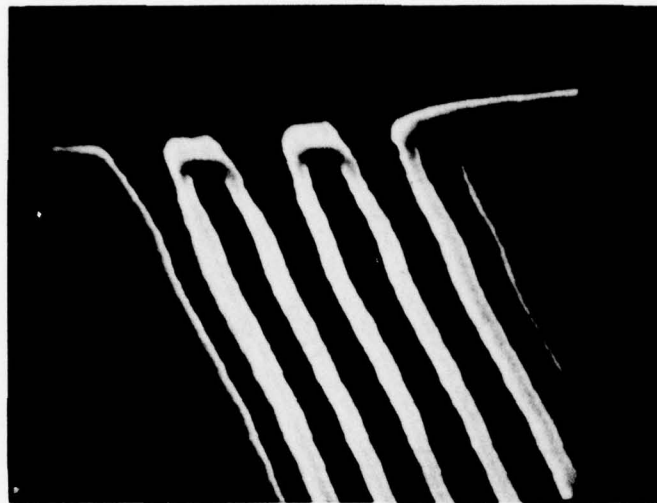
d = 0.3 μm F = 2.7 GHz

fig. 3   scanning electron micrograph of 0.3 μm wide fingers
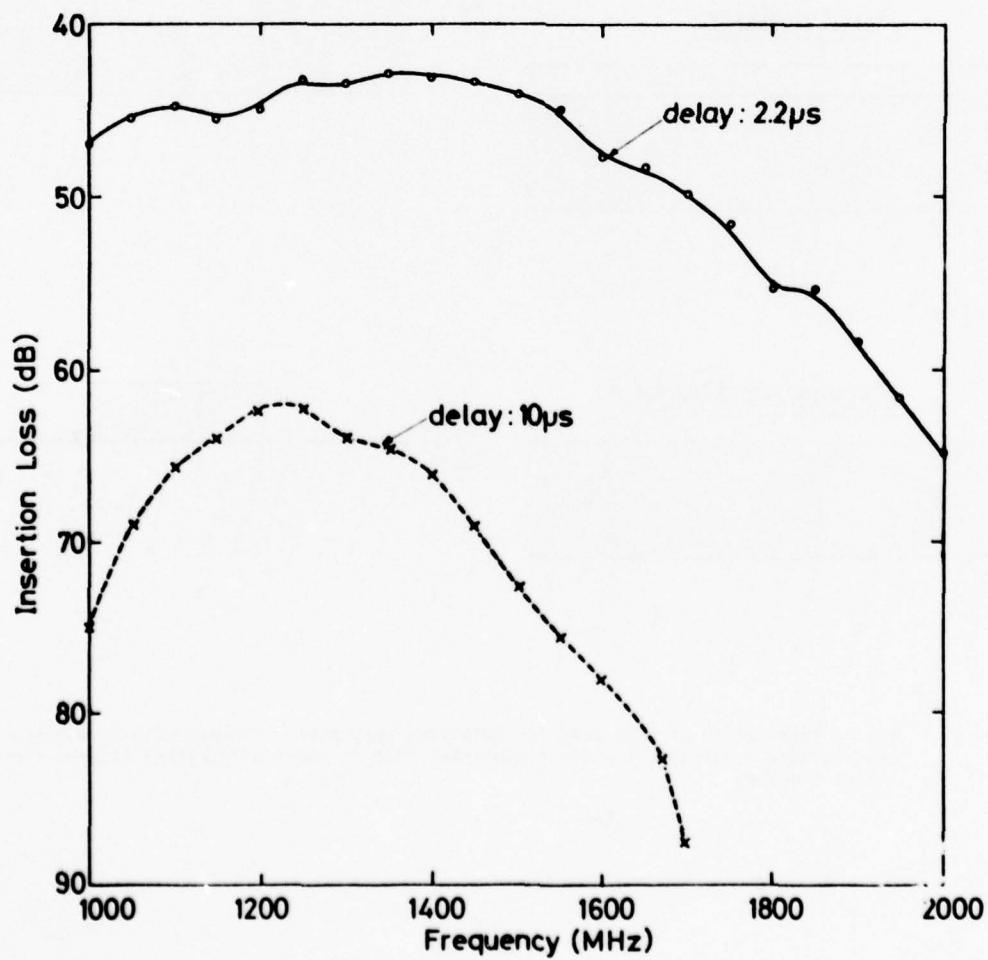
fig. 4   parts of frequency responses of two delay lines using (Y,Z) cut lithium niobate substrates.

output n°	delay μs	min. loss dB
1	1	26.6
2	2	28.5
3	3	30.5
4	4	33.3
5	5	37.8
6	6	40.1
7	7	42.1
8	8	44.9
9	9	47.9
10	10	52.6
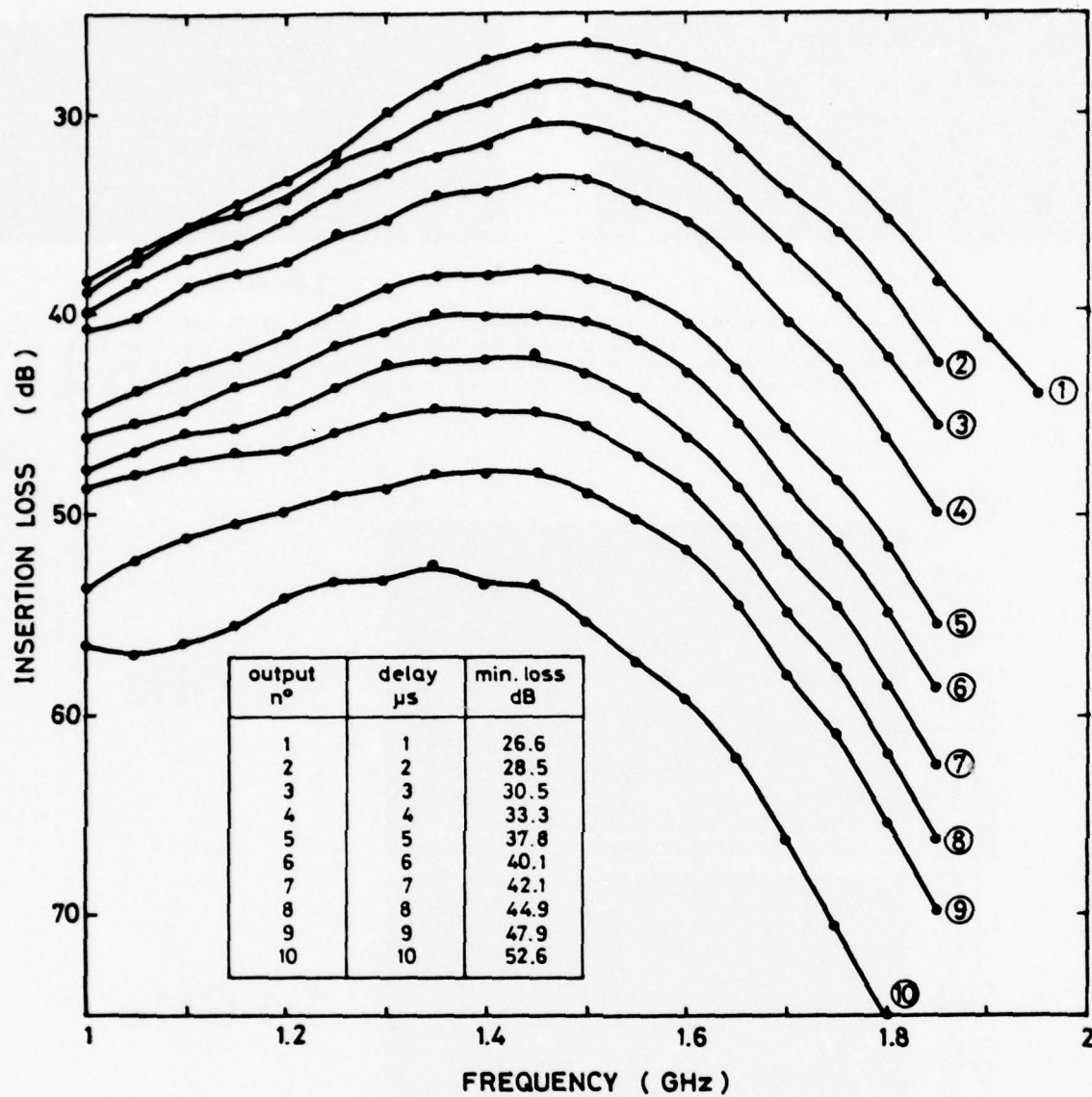
fig. 5   frequency responses of a tapped delay line using a (Y,Z) cut lithium niobate substrate.
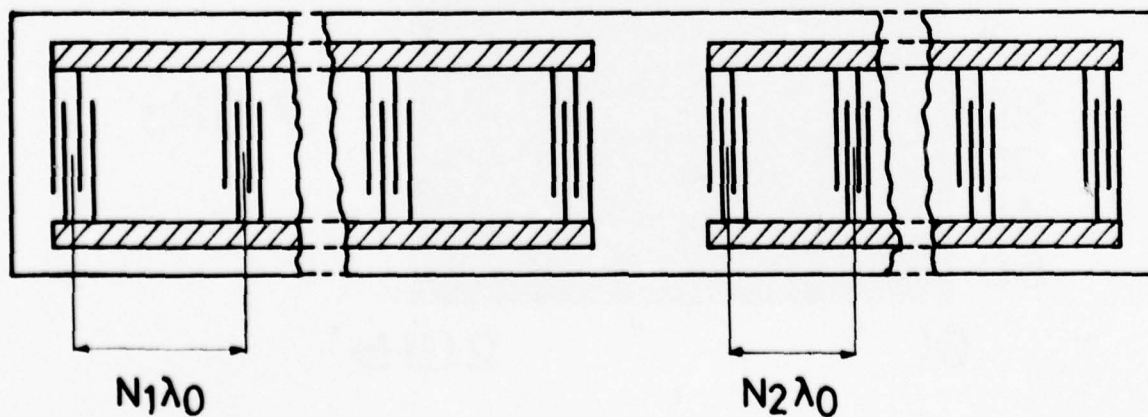
fig. 6   transducer configuration of a narrow bandwidth filter. N_1 and N_2 are two different integers and λ_0 is the acoustic wavelength at the centre frequency
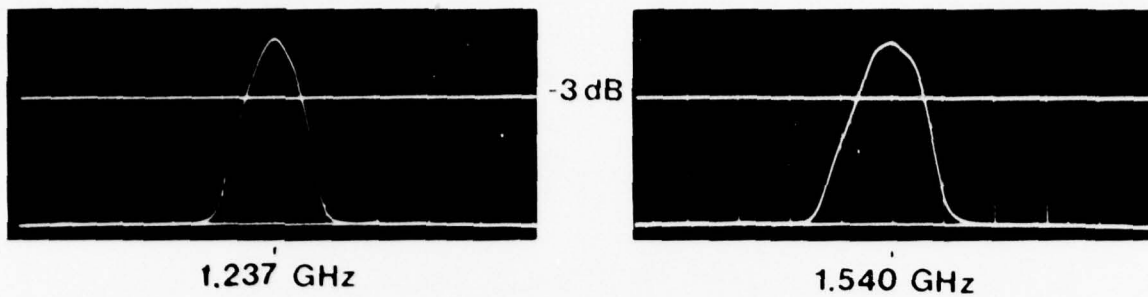
1.237 GHz   1.540 GHz

fig. 7   frequency response peaks of a narrow bandwidth filter using a (Y,Z) cut lithium
          niobate substrate. At 1.237 GHz the 3 dB bandwidth is 1 MHz and the insertion loss
          24.5 dB without impedance matching. At 1.540 GHz the 3 dB bandwidth is 1.2 MHz and
          the insertion loss 25.5dB. The side-lobe level of these peaks is inferior to 25 dB.

1.1 GHz

1.4 GHz

1.7 GHz
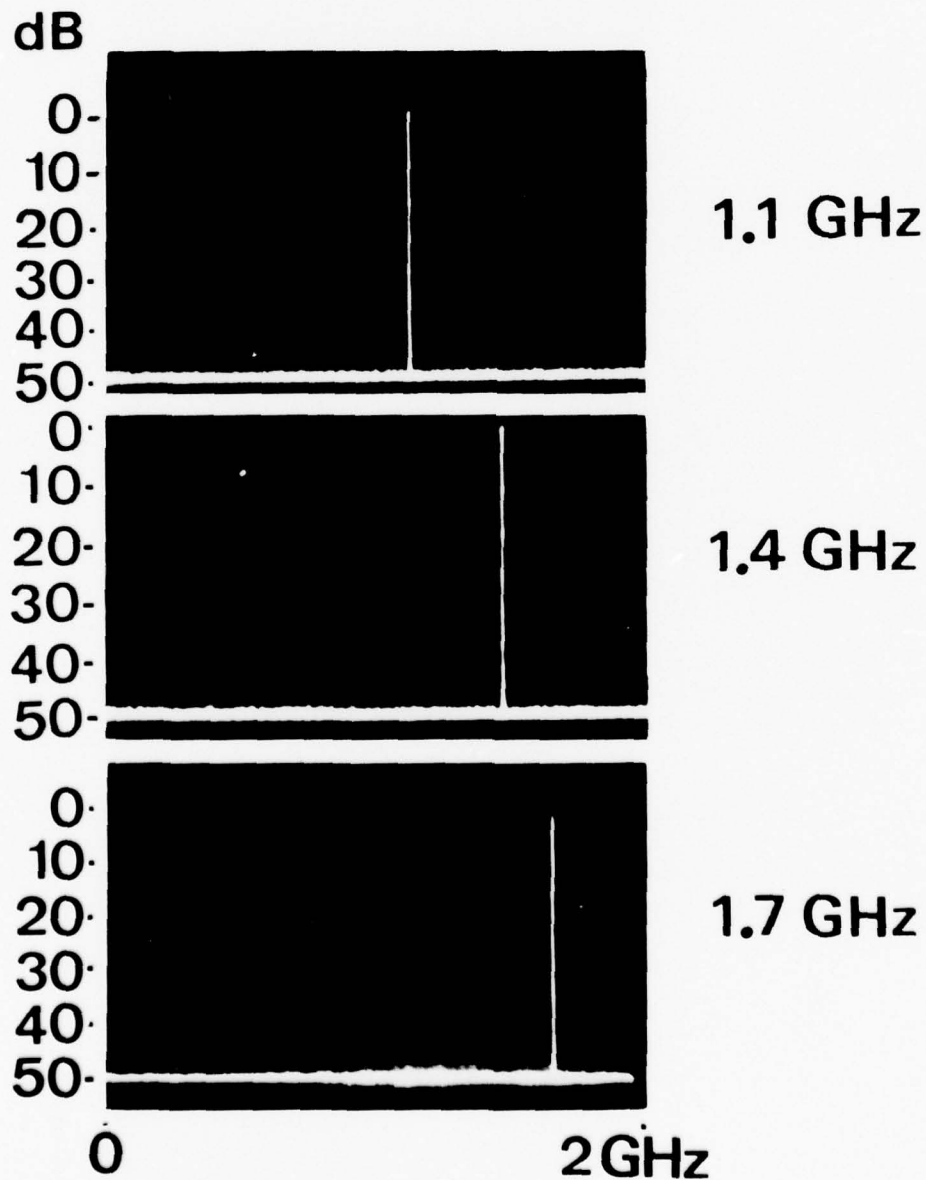
fig. 8   spectra of waves transmitted by a delay line oscillator using a ST cut quartz
          substrate. The 3 frequencies are obtained according to the frequency of impedance
          matching.
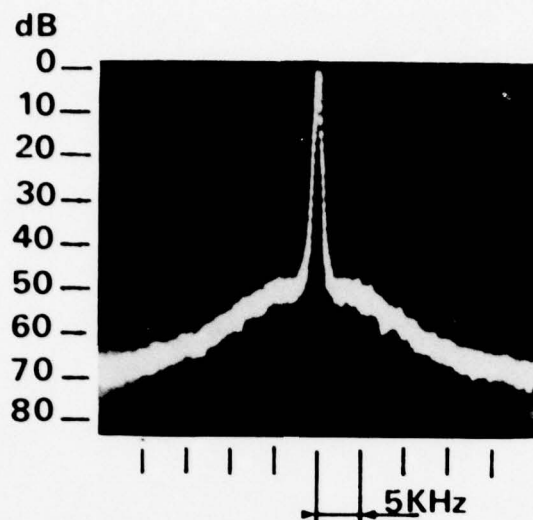
fig. 9   spectrum of 11 GHz oscillator implemented with a 1.01 GHz SAW delay line oscillator
The analysis bandwidth is equal to 300 HZ.

DISCUSSION

P.Carr

Is it possible to get 0.6 microns by using optical lithography?

Réponse d'Auteur

Par des procédés classiques de photo lithogravures il est possible d'obtenir des largeurs de doigts jusqu'à 3 à 4 microns. Le masqueur des électronique permet de descendre jusqu'à 0,3 microns et on peut l'utiliser dans la gamme 0,3 microns — 2 microns.

E.Stern

Could you contract the performance of your 11 GHz oscillators with a conventional microwave oscillator?

Réponse d'Auteur

Cet oscillateur a été réalisé par multiplication de fréquence à partir d'un oscillateur à 1 GHz. Il m'est difficile de répondre à votre question.

DEVELOPMENT OF A 100MHz BANDWIDTH PULSE COMPRESSION

SUBSYSTEM FOR AIRBORNE APPLICATION

M.B.N. Butler, B.Sc., PhD., M. Inst. P.,

Microwave and Electronic Systems Limited,
Lochend Industrial Estate,
Newbridge,
Midlothian,
EH28 8LP,
Scotland.

ABSTRACT.

A subsystem design incorporating S.A.W. devices for passive signal encoding and pulse compression with 100MHz bandwidth and 5μs dispersion is described. The paper is divided into three parts covering S.A.W. device design, subsystem design and environmental testing.

In order to minimise weight and power consumption the S.A.W. devices were designed for implementation on ST-X Quartz substrates. With this material and appropriate mechanical layout only minor changes in device performance are visible over an operating temperature range in excess of 100°C. In this development a new approach to the design of double-dispersive, angled chirp filters is taken. The technique has resulted in S.A.W. devices with T.B. = 500, insertion loss less than 45dB and matched to give a compressed pulse width of 15ns. with 34dB sidelobe suppression.

Subsystem design, including high gain, broadband amplifiers has shown the need for care in mechanical and electrical layout and screening. Attention is given to design for low R.F.I. and good pulse compression performance in a unit operating in an airborne environment.

The design and testing of the subsystem is illustrated with operational performance data and limitations observed during prototype test are outlined.

1. INTRODUCTION.

   The use of interdigital transducers on piezo-electric substrates was described by Smith et. al. (1969). The principles outlined by Smith have been widely described and developments into dispersive delay have formed a significant part of this work. Two substrate materials are commonly used. ST-X Quartz has a particular advantage over most others in that the temperature coefficient of delay for Rayleigh waves has a zero-point at normal ambient temperature. Quartz does, however, have a low electro-mechanical coupling coefficient and consequently high device loss must be accepted if large fractional bandwidths are required. Y-Z Lithium Niobate has a very much larger coupling coefficient but presents significant problems in spurious mode suppression, high array capacitance and poorer reproducibility than Quartz.

   After making the choice of ST-X Quartz for a chirp subsystem with large bandwidth a further compromise has to be made between the use of higher centre frequency with reduced fractional bandwidth or a lower centre frequency with an associated relaxation in electrode dimensions and thus improved processing uniformity. With the choice of 300MHz as the centre frequency it was accepted that 33% bandwidth would present high device insertion loss and poor impedance match into electronics employing normal circuit impedance of 50ohms. The associated requirement for high gain, wideband amplifiers has resulted in a modest transmitter signal-to-noise ratio of 20dB but this is found to be adequate since approximately 25dB of compression gain in the receiver results in source noise which is well below the sidelobe levels required and achieved.

2. S.A.W. DEVICE DESIGN.

   In the design of devices with 100MHz bandwidth at 300MHz it was decided that the type of structure shown in figure 1 should be employed. In this structure the two transducers are equally dispersive and thus have time-bandwidth products of 250. The inclined structure is employed to minimise complex, convolution effects and bulk wave coupling which can lead to degradation in the signal performance. Each device consists of an unapodised transducer and an apodised transducer; in the compressor spectral weighting is included to achieve sidelobe suppression. Each transducer is designed with interdigital electrodes with one series-connected break in each pair of electrodes. This structure has three particular advantages
   a) A larger aperture can be employed to reduce diffraction effects without creating arrays with excessive capacitance;
   b) The series connected structure is more tolerant of inter-electrode shorts since there are two gaps between the pads.
   c) Series connected electrodes can be employed in the angled structure without the danger of significant wavefront distortion since the breaks occur equally distributed across the aperture.

2.1. Multi-source design.

In a double ended device structure with angled transducers simple geometrical arguments show that the array over which accoustic convolution occurs changes in a complex fashion. In the case of a weighted compressor device in which delay increases with frequency this situation is illustrated in figure 1. In this figure it is assumed that the transducer to the left is unapodised and to note the performance at low, mid and high frequencies the areas A, B and C have been emphasised. As is customary in devices employing this type of structure the length ℓ is shown to be close to $L/\sqrt{TB}$, i.e.

$$\ell = 0.0632 \, L,$$

which results in signal contributions from all electrodes within the length ℓ. Clearly the amplitude function will be distorted by the asymmetry of the active region at all frequencies away from mid-band.

In the discussion above it was assumed that the unapodised transducer generates a plane wavefront and that this is retained as the wave propagates over a distance which is 23 times the source aperture at low frequencies and 55 times the source aperture at high frequencies. Over such distances significant diffraction spreading of the wavefront occurs. The significance of this effect is illustrated in figure 2 where the apertures and diffraction effects on the wavefront are shown approximately to scale. At all frequencies conditions approaching the far-field or point source circumstance apply at the exactly matched area. Perhaps a more important effect is the translation of energy laterally from the geometrical image condition shown in figure 1 to areas of the transducer outside that area. This lateral spread is quantified by noting that the 3dB width of the beam at the receiving transducer increases from 1.1a at low frequencies to 3.7a at high frequencies, where a is the transmitting aperture. Over the same range the peak amplitude decreased by a factor 0.57.

2.2. Result of implementing a multi-source design.

A mathematical derivation of the multi-source design is presented in the appendix. The main effect of this change in design approach is seen in figure 3. If the apodisation were implemented only on the basis of combining frequency weighting with the f^3 function derived through taking account of finger density effects (Hartmann (1973)) the apodisation function would be represented by the solid line. With attention to diffraction through application of a single source model the apodisation function is shown by the dot-dashed line and with the present multi-source model which extends the diffraction modelling and also takes account of acoustic convolution the apodisation function is shown by the dashed line. When the three functions are normalised to the value at mid-point in time one feature is particularly clear, the multi-source model introduces a correction which is in the opposite direction to the single-source model! A Fourier analysis of this error based on the paired echo theory (Klauder (1960)) indicates that the difference between the two models would produce a pair of sidelobes at time position 1/B having peak amplitude only 20dB below the main pulse. Such a pair of sidelobes would give significant pulse distortion and additional sidelobes might be expected.

3. SUBSYSTEM DESIGN.

The circuits employed for the passive expansion process are shown in figure 4. As a result of general design constraints the unit had to provide synchronisation of the trigger pulse to an externally available 300MHz COHO. After synchronisation a high level impulse was generated by causing a transistor to avalanche when triggered by the interface circuitry. Following the expansion process the amplifier was split into a low noise preamplifier stage and the main amplifier prior to gating a limiting. In general very little limiting was employed since this helps to maintain signal to noise and it was found that the S.A.W. impulse responses was flat enough to meet specification.

A feature of the expansion unit design was the requirement for circuit screening and earth contact. To achieve the performance required each circuit was implemented in circuit boards with a large area earth plane on the reverse side. These circuits were then inverted into pockets in a milled aluminium, module box with each amplifier package grounded through a heat sink. This configuration was found to minimise spurious oscillation and feed-through effects.

The compressor unit block diagram is shown in figure 5. In most respects the circuit implementation was straightforward and employed part of the same module box as the expansion unit. A feature of some concern was the receiver performance in the presence of many, moderate level returns: a situation which can occur quite readily in a system employing a time-bandwidth product of 500. To achieve suitable performance in this respect the amplifier characteristics shown in figure 6 were employed. In this the preamplifier exhibits little sign of saturation at total input levels in excess of +5dBm although the normal signal level for maximum output is -10dBm. In this way something in excess of 20 overlapping, high level returns can be processed simultaneously without significant precompression intermodulation or post compression limiting. Clearly the provision of such "headroom" reduces the dynamic range available below a single return of normal, maximum amplitude. This compromise was considered essential for system operation.

4. SYSTEM PERFORMANCE.

System performance falls into two general areas,
a) basic signal characteristics at normal ambient, and
b) variation of these characteristics in extreme environments.

4.1. Basic signal characteristics.

The expander unit output response is shown in figures 7 and 8. Figure 7 shows the impulse response and demonstrates the high standard of amplitude flatness achieved for an S.A.W. device operating at 300MHz. In figure 8 the expansion unit output spectrum demonstrates the moderate signal to noise ratio of 20dB. In the example available when these records were obtained the spectral amplitude was found to vary over a range of ±1.5dB. Since this variation exceeds the amplitude variations in the impulse response it is felt that this indicates phase variation which can probably be traced to variations in electrode width to space ratio at some places in the transducer array.

Figure 9 shows a compressed pulse output from the subsystem with a single input at -10dBm. In this example the -3dB pulse width is 15ns and receiver compression gain has improved the signal to noise ratio to an extent where sidelobes at -34dB are clearly visible. Detailed study of the timing synchronisation was achieved by phase detection of the compressed pulse. This test showed that phase jitter was below the measurement limit of ±3° which corresponds to a timing accuracy of ±28psecs.

4.2. Environmental performance.

4.2.1. Temperature.

Normal operation was required for -5° to +60°C with power-on warm up from -40°C and +70°C. In practice very little change in the basic performance was observed over the extended temperature range. The main problem was found in gain variations with temperature which amounted to ±1dB in the expansion unit and a further ±1dB in the compressor unit.

Sidelobe levels were particularly good at high temperature with -34dB being the worst observed from +20°C to +70°C. At lower temperature some degradation occured with a single sidelobe increasing to -30dB at -5°C and to -28dB at -40°C. This effect only appeared temporarily and improved after 20 minutes operation at -5°C when turned on from cold soak at -40°C.

4.2.2. Acceleration and vibration.

Normal performance was observed during acceleration to 10g in each of three orthogonal axes and no damage was observed during electrical and visual testing after vibration to 1.5g peak over the frequency range 15Hz to 500Hz.

4.2.3. Spurious injection on the power supplies.

The unit employs supplies at +24 Volts, +15 Volts and -5.2 Volts. In each case tests were carried out with injected spurious at 0.3 Volts peak to peak over the frequency range 400Hz to 500MHz. During tests on the 24 Volt and 15 Volt supplies slight phasing effects were observed on the sidelobes but at no time did the effects result in more than 1dB change in level. Some effect was observed in timing jitter but this was always less than the ±1° observed with high frequency injection on the -5.2 Volt line. The latter sensitivity was considered unimportant in practice since this power line was only used to drive ECL circuits which would tend to carry only trigger synchronous ripple.

The results of performance evaluation are summarised in the tables below.

Expander:	Specification	Observed
Centre frequency	300 ± 3MHz	300.3MHz
Bandwidth	100 ± $_0$MHz	100.4MHz
Dispersion	5.0 ± 0.1μs	4.98μs.

Compressor:	Specification	Observed
Compressed pulse width	16ns max.	16ns worst case.
Sidelobe level	-30dB peak max.	-34dB above R.T. -30dB worst case at -5°C.
Pulse jitter	3° max.	<< 3° max.

5. CONCLUSIONS.

A high frequency, wideband pulse compression subsystem with a time-bandwidth product of 500 has been developed. The unit employs S.A.W. devices which were designed through application of novel design procedures which have proved to be very successful. Although neither the S.A.W. devices nor the electronics were temperature controlled performance has been shown to be satisfactory from -40°C to $+70^{\circ}$C. Detailed evaluation under conditions appropriate to a real, airborne environment have shown that the subsystem is unconditionally stable and maintains a high order of pulse compression performance and timing coherence under all circumstances.

6. ACKNOWLEDGEMENTS.

7. APPENDIX.

When a broadband impulse is applied to a double dispersive transducer structure the output signal can be expressed as the convolution of the impulse responses of the two individual transducers. The integral, $h(t)$, represents the response when the signal transmitted by the constant aperture transducer P passes under and is detected by the weighted transducer Q. If the impulse responses of P and Q alone are given by $p(t)$ and $q(t)$, respectively then:

$$h(t) = \int p(t-\tau) \, q(\tau) \, d\tau.$$

When P and Q represent linear chirps we have:

$$p(t) = \exp j(\omega_o \tau + \mu\tau^2/2)$$

$$q(t) = w(\tau) \exp j(\omega_o \tau + \mu\tau^2/2); \quad -\frac{T}{2} < \tau < \frac{T}{2}$$

i.e.
$$h(t) = \exp j(\omega_o t + \mu t^2/2) \int_{\tau_1}^{\tau_2} w(\tau) \exp j(\mu\tau^2 - \mu\tau t) \, d\tau \tag{1}$$

where $-T<t<0$; $\tau_2 = \frac{T}{2} + t$; $\tau_1 = -\frac{T}{2}$.

$0<t<T$; $\tau_2 = \frac{T}{2}$; $\tau_1 = t-\frac{T}{2}$.

This complex convolution integral is evaluated numerically using a small enough step to describe accurately the variation of $\exp j(\mu\tau - \mu\tau t)$ with τ.

The two dimensional nature of surface wave propogation from an interdigital transducer is analagous to the optical case of light illuminating a thin slit of length 2a. It is assumed that the transducer is represented as a line source with constant amplitude per unit length. If propogation takes place along the x-axis, and provided $x > 2a$, the amplitude of the acoustic beam at an observation point (x,y) is given by:

$$A(x,y) = 2^{-\frac{1}{2}} \int_{y^-}^{y^+} \exp(-j\pi v^2/2) dv \tag{2}$$

where
$$y^+ = (y+a) / (\lambda x/2)^{\frac{1}{2}}$$

$$y^- = (y-a) / (\lambda x/2)^{\frac{1}{2}}$$

If the receiving electrode is of length 2b then the total detected signal amplitude $A(x)$ is given by:

$$A(x) = 2^{-\frac{1}{2}} \int_{-b}^{b} \int_{y^-}^{y^+} \exp(-j\pi v^2/2) \, dv. \, dy. \tag{3}$$

Normalising to the acoustic wavelength, λ, (2) may be rewritten as:

$$A(x) = 2^{-\frac{1}{2}} \int_{-\hat{b}}^{\hat{b}} \int_{\hat{y}^-}^{\hat{y}^+} \exp(-j\pi v^2/2) \, dv. \, dy. \tag{3a}$$

where $\hat{b} = b/\lambda$; $\hat{a} = a/\lambda$ and $\hat{x} = x/\lambda$

The above theory applies to isotropic substrates and a modification must be made to take account of the anisotropy of ST-X cut Quartz. It has been shown by Szabo (1973) that a useful theory for calculating anisotropic diffraction fields is to assume that on or near the axis of propagation velocity can be approximated by a parabolic law. Cohen (1967) has compared a rigorous analysis of anisotropic diffraction with the approximate analysis. He concludes that anisotropic diffraction can be represented by the Fresnel integral of equation (1) with a scale change.

$$x^1 = x(1 + \gamma) \tag{4}$$

where γ is an anisotropy factor dependent on the square law velocity variation in the vicinity of the pure made axis defined by the velocity equation:

$$v(\theta) = v_o (1 + \gamma\theta^2/2) \tag{5}$$

The value of γ for ST-X cut Quartz has been computed by Campbell and Jones (1968) from velocity data : they find $\gamma = 0.378$.

If x in (3a) is appropriately scaled according to (4), one obtains the simplified expression:

$$A(x) = \left(\frac{\hat{x}}{2}\right)^{\frac{1}{2}} \left[\alpha_+ F(\alpha_+) - \alpha_- F(\alpha_-) - \frac{j}{\pi} \left\{\exp(-j\pi\alpha_+^2/2) - \exp(-j\pi\alpha_-^2/2)\right\}\right] \tag{6}$$

where $F(\alpha) \int_o^\alpha \exp(-j\pi v^2/2)$ is the Fresnel integral and

$$\alpha_\pm = (\hat{b} + \hat{a}) / (\hat{x}/2)^{\frac{1}{2}} \tag{7}$$

Equation (6) describes the situation when both transmitter and receiver lie along the same axis. If there is a lateral separation c between transmitter and receiver as in the inclined transducer structure, then the received voltage is given by:

$$A(\hat{x},\hat{c}) = 2^{-\frac{1}{2}} \int_{-\hat{b}+\hat{c}}^{\hat{b}+\hat{c}} \int_{y-}^{y+} \exp(-j\pi v^2/2) \, dv. \, dy. \tag{8}$$

Simplification of this expression leads to:

$$A(\hat{x},\hat{c}) = \left(\frac{\hat{x}}{2}\right)^{\frac{1}{2}} \left[\alpha F(\alpha_1) + \alpha_2 F(\alpha_2) - \beta_1 F(\beta_1) - \beta_2 F(\beta_2)\right.$$

$$\left. -\frac{1}{\pi} \left\{\exp(-j\pi\alpha_1^2/2) + \exp(-j\pi\alpha_2^2/2) - \exp(-j\pi\beta_1^2/2) - \exp(-j\pi\beta_2^2/2)\right\}\right] \tag{9}$$

where

$$\alpha_1 = (\hat{b} + \hat{a} + \hat{c}) / (\hat{x}/2)^{\frac{1}{2}}$$

$$\alpha_2 = (-\hat{b} - \hat{a} + \hat{c}) / (\hat{x}/2)^{\frac{1}{2}}$$

$$\beta_1 = (\hat{b} - \hat{a} + \hat{c}) / (\hat{x}/2)^{\frac{1}{2}}$$

$$\beta_2 = (-\hat{b} + \hat{a} + \hat{c}) / (\hat{x}/2)^{\frac{1}{2}}$$

The procedure for synthesising the double inclined transducer structure consists of the evaluation of equation (1) subject to the constraint that the sample weights $w(\tau)$ are evaluated by using the precise diffraction integral of (9). If $A(\hat{x},\hat{c})$ replaces $w(\tau)$ in (1) then the full expression for the received signal impulse response is given by:

$$h(\tau) = \int_{\tau_1}^{\tau_2} A(\hat{x},\hat{c}) \exp\left[j\mu(t2/2 + \tau^2 - \mu\tau t)\right] \, d\tau.$$

where

$$\hat{x} = v_s t (1 + \gamma) / \lambda$$

$$\hat{c} = v_s \tau \tan\theta / \lambda$$

and θ is the angle of inclination of the transducer.

8. UNDERLINE REFERENCES.

Campbell, J. J., Jones, W. R., (1968), IEEE Trans SV-15, 1223.

Cohen, M. G., (1967), Journal of Applied Physics, p. 38.

Hartmann, C. S., Bell, D. T., Rosenfeld, R. C., (1973), IEEE Trans MTT-21, 162.

Klauder, J. F., Price, A. C., Darlington, S., Albersheim, W. J., (1960), Bell System Tech. Journal 34, 745.

Smith, W. R., Gerard, H. M., Collins, J. H., Reeder, T. M., Shaw, H. J., (1969), IEEE Trans MTT-17, 856.

Szabo, T. L., Slobodnik, A. J., (1973), IEEE Trans SU-20, 240.

SERIES
CONNECTED
ELECTRODES

LOW
FREQUENCY

A — CONVOLUTION
AREAS

B

HIGH

FREQUENCY

C

L

TRANSMITTER
(UNAPODISED)

RECEIVER
(APODISED)
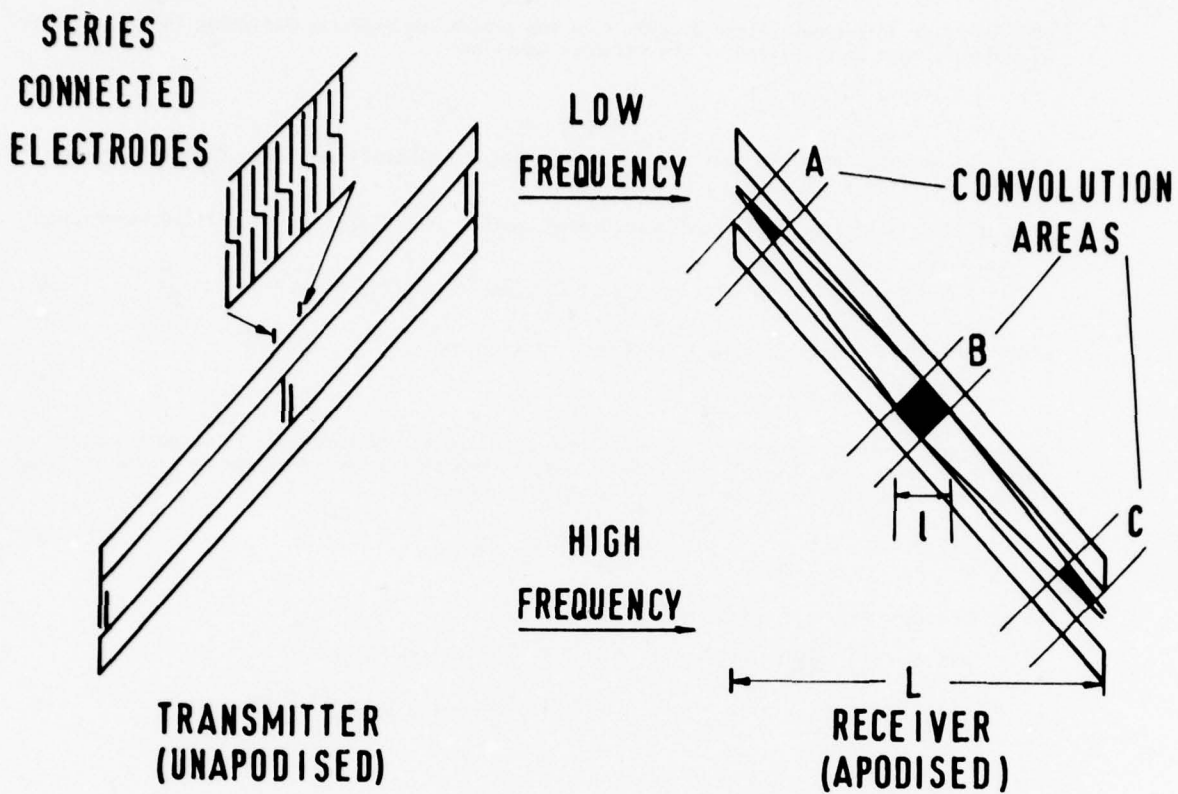
Figure 1.   Double dispersive inclined transducer geometry employing series
connected electrodes.

LOW FREQUENCY

TRANSMITTER

RECEIVER
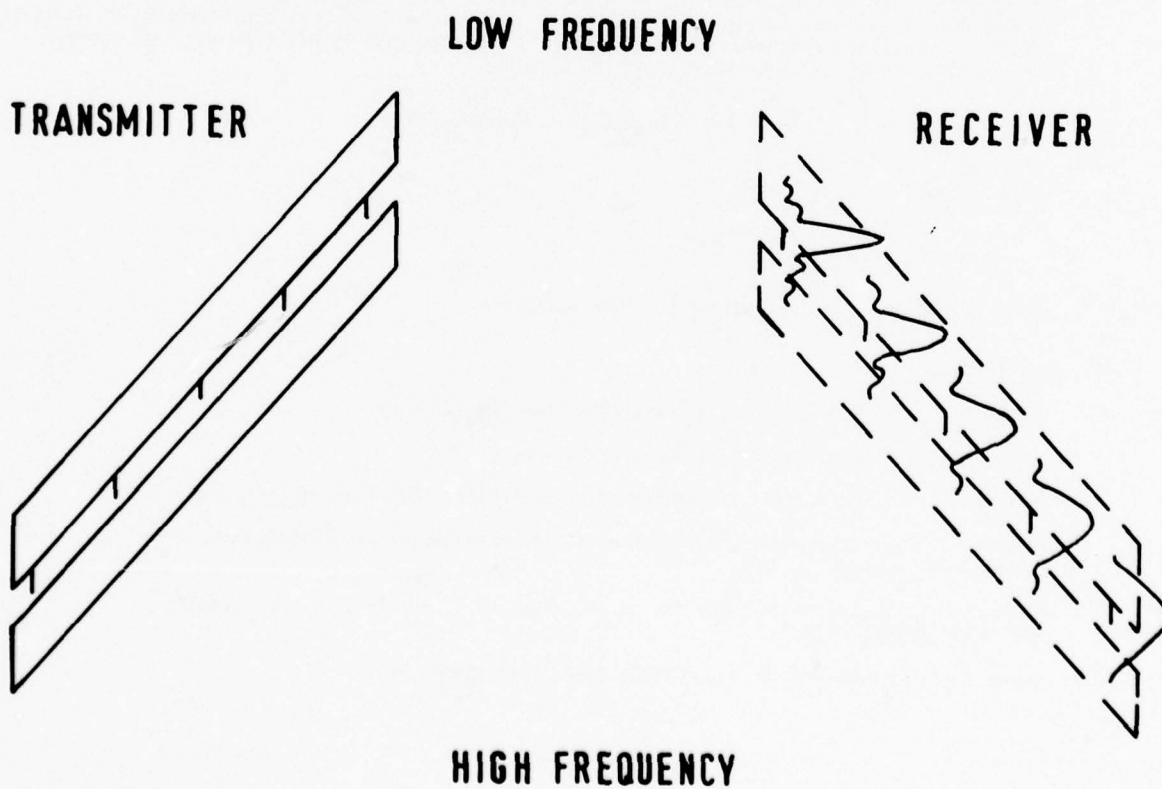
HIGH FREQUENCY

Figure 2.   Diffraction effects in high frequency inclined transducer geometry.
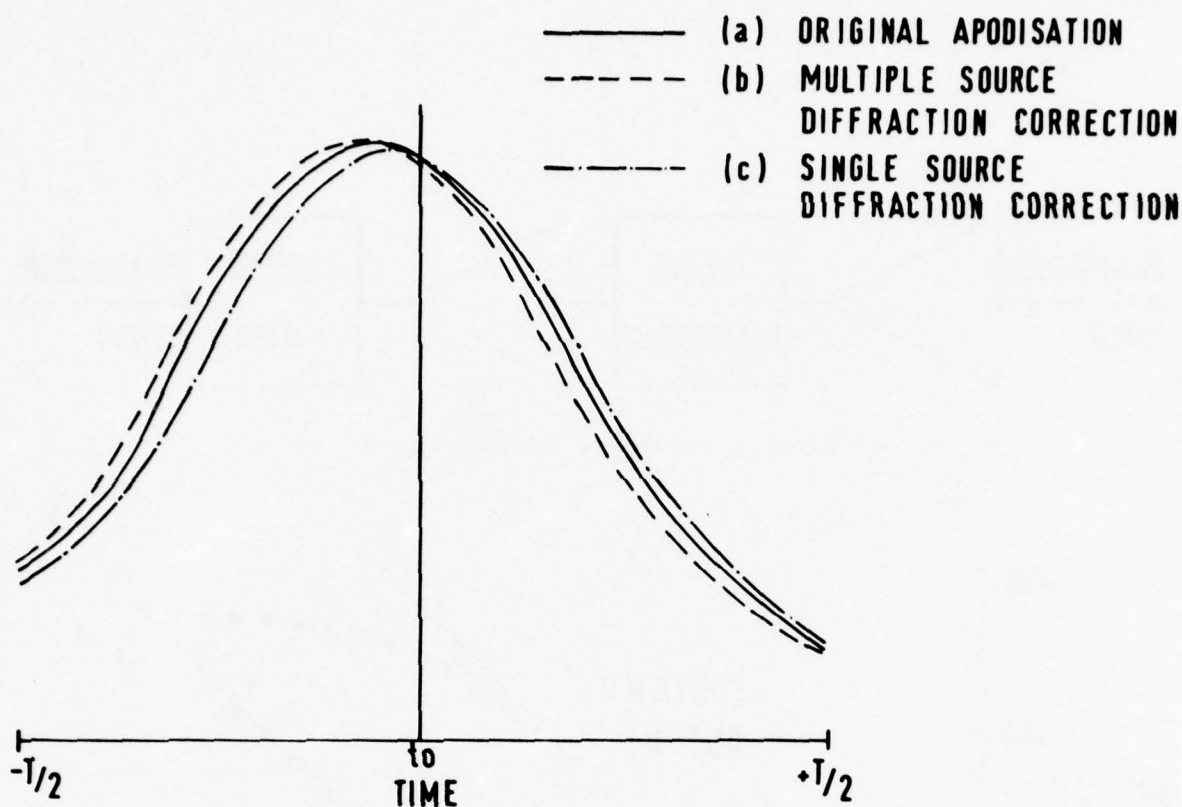
Figure 3. Apodisation profiles for various models with and without diffraction correction.
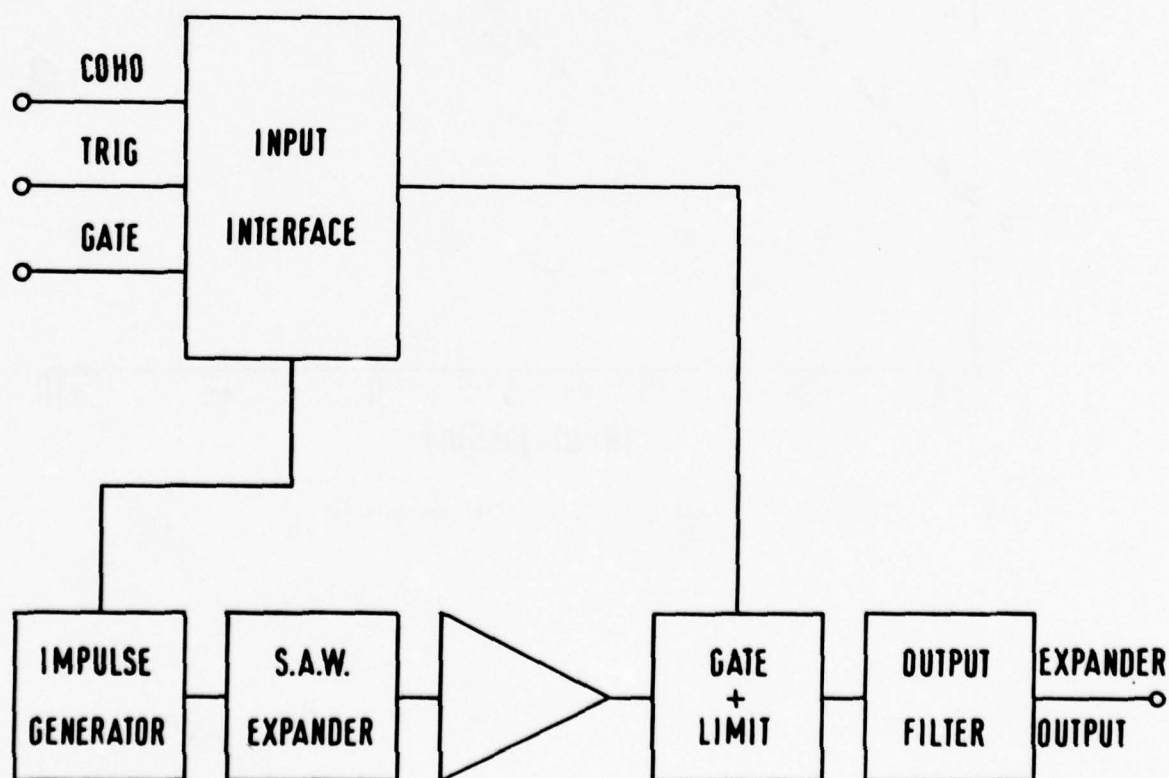
Figure 4. Expansion unit block diagram.
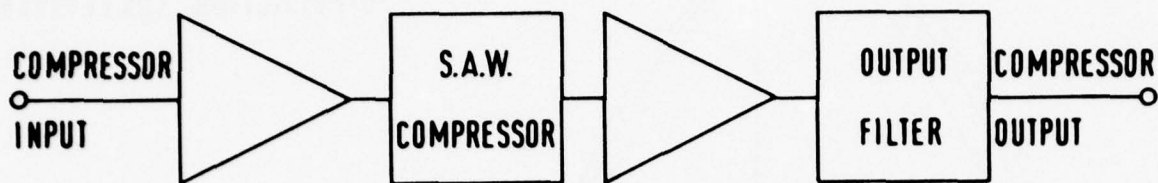
Figure 5.   Compressor unit block diagram.
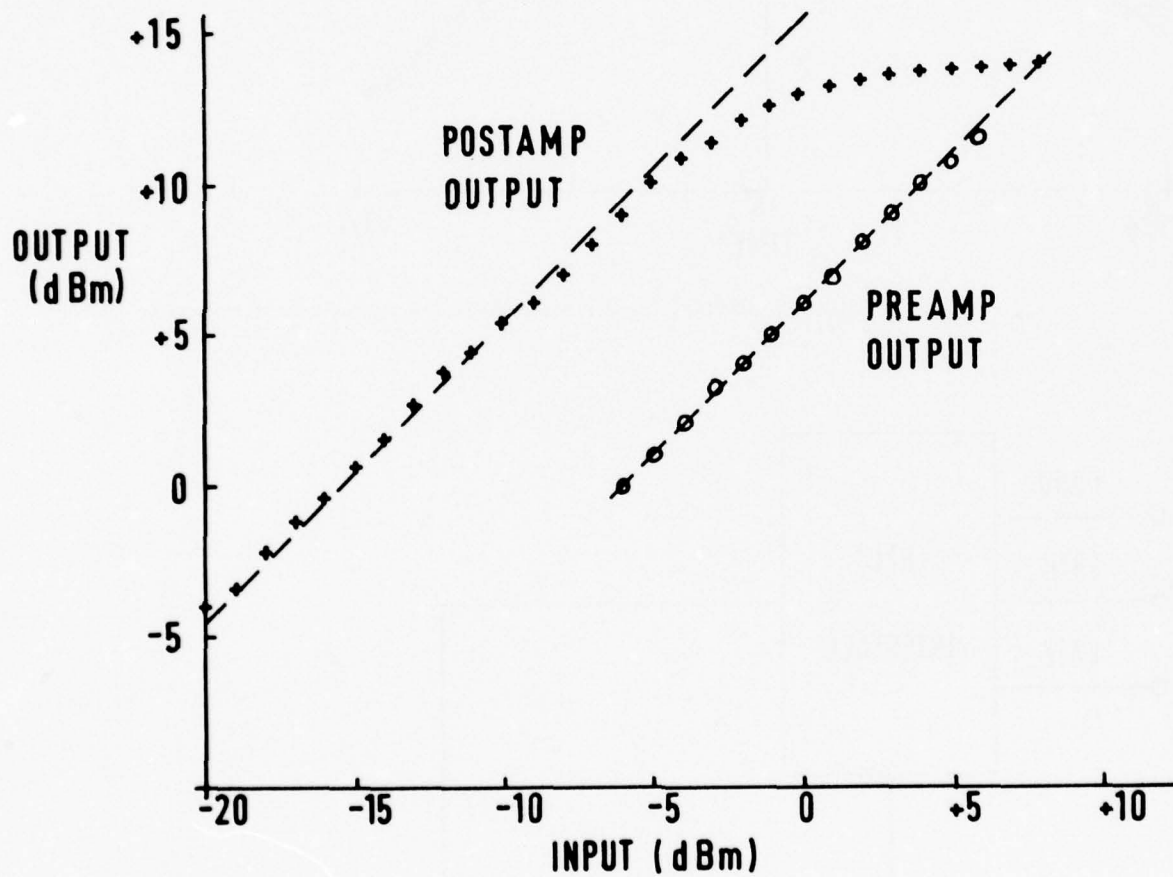
Figure 6.   Compressor unit amplifier characteristics.
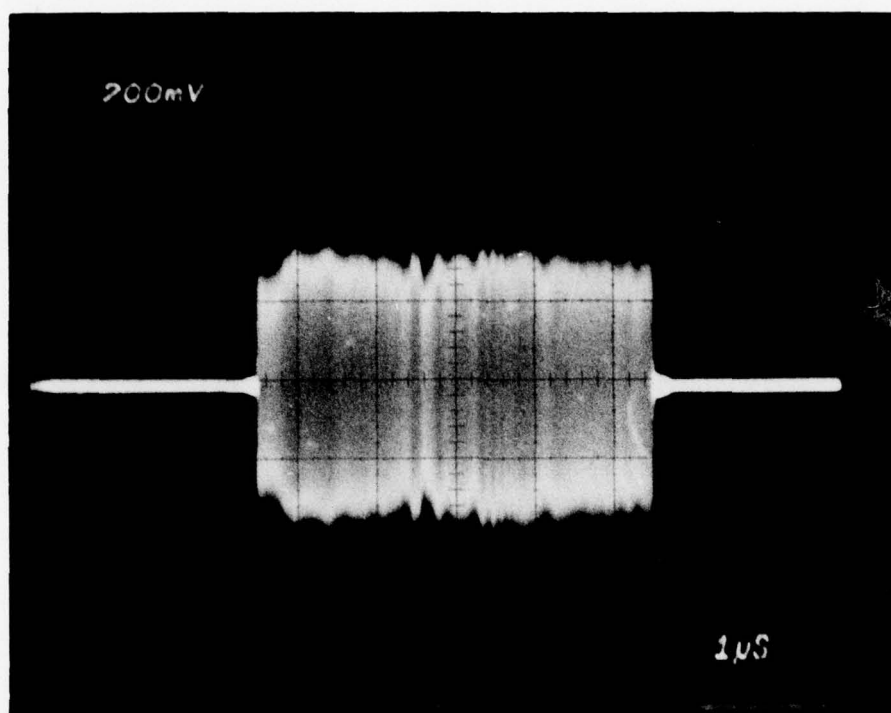
Figure 7. Expansion unit impulse response.
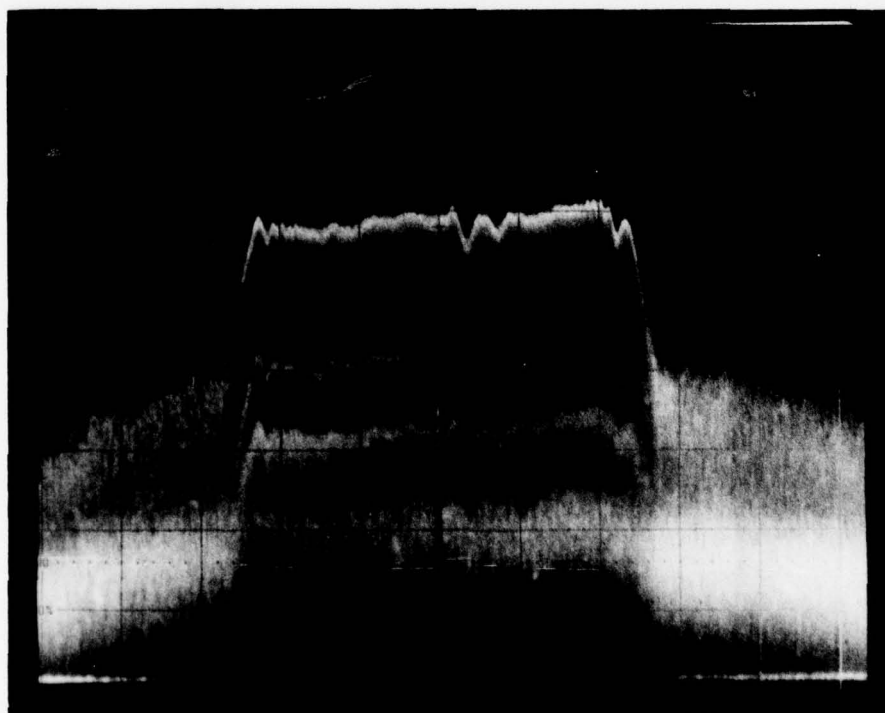
Figure 8. Expansion unit spectrum.
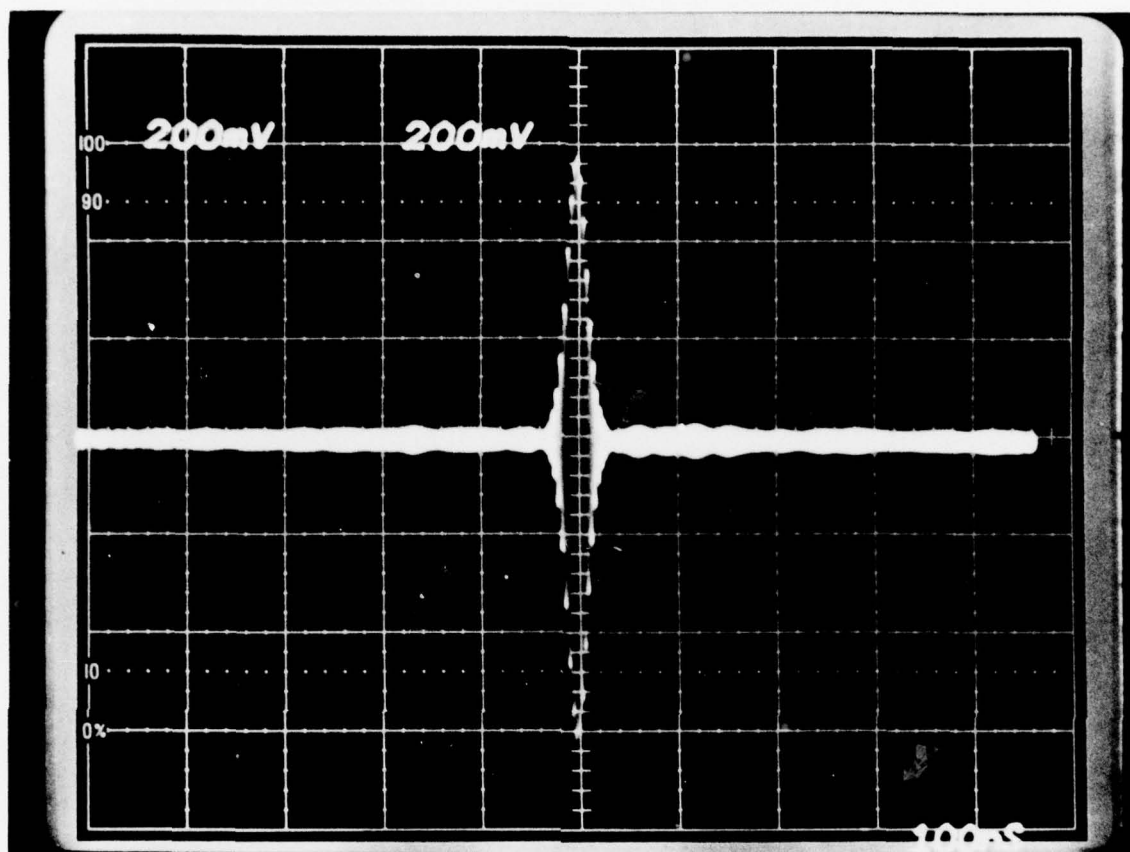
Figure 9.   Compressed pulse.

DISCUSSION

E.Stern

What was the S/N ratio of expander output?

Author's Reply

Expander output S/N ratio was 20 db. This is significantly poorer than should have been obtained with a device loss of 45 db in a 50-ohm system. A possible cause is the mismatch of the avalanche impulser and the SAW device input.

SIGNAL PROCESSING WITH A REFLECTIVE DOT ARRAY (RDA)

L. P. Solie and H. van de Vaart
Sperry Research Center
Sudbury, Massachusetts 01776

SUMMARY

A new type of pulse compression filter is described using the Reflective Dot Array (RDA). The RDA is similar to the Reflective Array Compressor (RAC), except that the array of reflecting grooves is replaced by an array of reflecting metallic dots. The RDA has the principal advantage of being part of the same mask and metalization as the interdigital transducers, allowing single-step fabrication. A linear FM filter was developed with a center frequency of 60 MHz, bandwidth of 20 MHz and differential time delay of 10 μs, with less than 3° of rms phase deviation from quadratic without phase compensating film, showing that high performance pulse compression filters can be produced at low cost.

1. INTRODUCTION

The reflection of surface acoustic waves by an array of grooves has been used for several years to fabricate high performance pulse expansion and compression filters, as well as bandpass filters (WILLIAMSON, R.C. and H. I. SMITH, 1973) and (DOLAT, V. and J. MELNGAILIS, 1974). The advantages are well known: 1) bulk modes and pseudo surface modes are not detected by the output transducers because they are reflected at a different angle; 2) the structure is defect tolerant; 3) there is no regeneration; 4) grooves are separated by $\lambda\sqrt{2}$, whereas electrodes with split fingers are separated by $\lambda/4$, the magnitude of the reflection is only dependent on the depth of the groove and not on the piezoelectric coupling, thereby increasing the flexibility of achieving the desired tap weight; 6) the depth weighting provides true amplitude weighting over the full beam width, eliminating diffraction problems and 7) twice the differential time delay can be obtained for a given length substrate. The major disadvantages lie in the fabrication process: 1) a complex ion beam etching system is necessary to provide the variable depth grooves; 2) the process is serial, so that devices can be fabricated only one at a time and 3) precise alignment is necessary between the metalized transducers (produced in one step) and the array of grooves (produced in another step).

In previous papers, it was shown that the reflections of surface waves can be accomplished by an array of metalized dots (RDA), where the number of dots in each row determines the strength of the reflection, i.e., provides amplitude weighting (SOLIE, L.P., 1976) and (SOLIE, L.P., 1976). This principle was applied to the construction of a high performance bandpass filter. The advantages of using the RDA are that in addition to all the advantages of the reflective array compressor as listed above, the amplitude weighting is built into the same mask that contains the transducers; both transducers and the array of dots are evaporated into the substrate at the same time, eliminating the critical alignment step, and this process allows fabrication of many devices in parallel.

In this paper, we report on the design of pulse compression filters using the RDA approach. The design procedure is in many respects identical to the RAC device. However, there is a difference in the manner in which the phase compensation due to the mass loading of the metal dots is calculated as compared to the grooved devices. The design of a pulse compression filter with a time-bandwidth product of 200 and the results obtained are described in the next sections.

2.   RDA DESIGN

In analogy to the transfer function for grooved devices (WILLIAMSON, R.C. and H. I. SMITH, 1973) the transfer function for an RDA can be written as

$$H(\omega) = \left(\frac{Ch\omega}{2\pi v}\right)^2 \sin^2\left(\frac{\omega d}{2v}\right) \sum_{m,n=1}^{N} a_m a_n \lambda_{mn} e^{-j\frac{\omega}{v}(x_m + x_n)}$$ (1)

where C is a proportionality constant relating the reflectivity to the height of the discontinuity (here the thickness h of the metal dot), v is the surface wave velocity, d is the dimension of the dot in the direction of propagation (assumed constant over the array), a_m and a_n represent the normalized weighting for each row in the two arrays, λ_{mn} is the overlap function which measures the fraction of the n-th row illuminated by the m-th row, and x_m and x_n measure the distance from the first row to the m-th and n-th rows, respectively. As will be shown in the following section, using Eq. (1) predicts the transfer characteristics of the RDA device with good accuracy.

In calculating the position of each row, one has to consider the velocity reduction which is a function of the density and size of the dots, the metal thickness and the frequency. The fractional velocity reduction is assumed to take the form

$$\frac{\Delta v}{v} = -KA\left(\frac{h}{\lambda}\right)^2$$ (2)

where K is a constant, A is the fractional area metalized and λ is the wavelength of the surface wave. This assumes that there is no continuous metalization, i.e., the dots are smaller than a wavelength so that there is no piezoelectric shorting. Equation (2) can be rewritten as

$$\frac{\Delta v_{ij}}{v_o} = \frac{-KNa_i}{16M}\left(\frac{h}{\lambda_o}\right)^2\left(\frac{f_j}{f_o}\right)^2\left(\frac{f_i}{f_o}\right)$$ (3)

where $\Delta v_{ij}/v$ is the fractional reduction of the phase velocity of a propagating wave of frequency f_j in the vicinity of a row at which frequency f_i is resonant, f_o is the center frequency and λ_o its corresponding wavelength, and $M\lambda_o$ is the width of the array. Equation (3) assumes the dots have an area of $(\lambda_o/4)^2$.

The group velocity for a wave at frequency f_j is given by

$$v_g = 2\pi\left(\frac{dk'}{df_j}\right)^{-1} \tag{4}$$

where the wavevector $k' = k(1+\frac{\Delta v}{v})^{-1}$. The group delay for a surface wave at frequency f_j up to the row at which f_j is resonant (located at x_j) is then found to be

$$\tau_g(f_j) = \frac{x_j-x_1}{v_o} - 3\sum_{i=1}^{j-1}\frac{\lambda_i}{v_o}\frac{\Delta v_{ij}}{v_o} \tag{5}$$

By fitting this expression to the required chirp slope, the positions x_j (relative to the position of the first row x_1) can be determined. These equations can also be applied to the design of an RDA bandpass filter, for which case $f_i = f_o$.

3. EXPERIMENTAL RESULTS

A linear FM filter was designed at a center frequency of 60 MHz, a bandwidth of 20 MHz and a differential time delay of 10 μs, for a time-bandwidth product of 200. The filter was overdesigned to 28 MHz bandwidth and 14 μs differential time delay in order to be able to gate out the ripple at the edges of the passband. Also, in order to prevent refraction of the surface wave as it enters the dot array, the edges of the array were cut off perpendicular to the propagation direction. This, of course, causes additional weighting at the edges of the passband, but helps to reduce the ripple. The input and output transducers were phase reversed transducers (9 fingers) with a measured insertion loss of 35 dB.

The computed passband response using Eq. (1), as well as the experimentally measured response, is shown in Fig. 1. Although the measured curve has a somewhat larger dip in the center of the band, the overall agreement is excellent. Gold dots were used with a thickness of 4400 Å; subtracting the transducer loss, the reflection loss is found to be ~ 10 dB per array.

The phase response was also measured, and fitted to a quadratic curve. The deviations from quadratic are shown in Fig. 2. The rms phase deviation is $2.3°$, an excellent result considering that no phase compensating film was used in between the two arrays. This response implies that, with appropriate weighting, pulse compressors with time sidelobes down 35 dB or better can easily be obtained using the RDA. Work on weighted RDA pulse compressors, as well as on devices with larger time-bandwidth products, is presently in progress.

Finally, two identical devices were used in a pulse compression loop. One device was impulsed to obtain the expanded signal (Fig. 3(a)). After spectral inversion, this signal was gated and then compressed in the second RDA (Fig. 3(b)). An expanded display of the compressed signal is shown in Fig. 3(c).

4. CONCLUSION

In this paper, we have reported the initial results obtained with a linear FM pulse expansion and compression filter with only a modest time-bandwidth product using the RDA approach. We believe that this approach will find wide application in the fabrication of high performance pulse compression filters, primarily because of the ease of fabrication while maintaining the excellent characteristics of the reflective array compressor. With the modern fast-scanning pattern generators, mask fabrication has posed no problem. The somewhat higher cost of the mask (because of the many flashes) is more than offset by the elimination of the complex ion beam etching apparatus. For very long time delay devices, it may prove advantageous to use etched holes (or posts) rather than metalized dots, but even in that case the etch depth is uniform and can be done in a simple sputter etch system--in parallel fashion rather than serial. For these reasons, the RDA approach shows promise to fabricate high performance filters at low cost.

REFERENCES

DOLAT, V. and J. MELNGAILIS, 1974, "16 channel surface acoustic wave filter bank," _Ultrasonics Symposium Proceedings_, p. 756.

SOLIE, L.P., 1976, "Surface acoustic wave reflective dot array (RDA)," _Appl. Phys. Letters_, vol. 28, 420.

SOLIE, L.P., 1976, "A SAW filter using a reflective dot array (RDA)," _Ultrasonics Symposium Preceedings_, p. 309.

WILLIAMSON, R.C. and H. I. SMITH, 1973, "The use of surface-elastic-wave reflection gratings in large timeband pulse-compression filters," _IEEE Trans. on Microwave Theory and Techniques_, vol. MTT-21, 195.
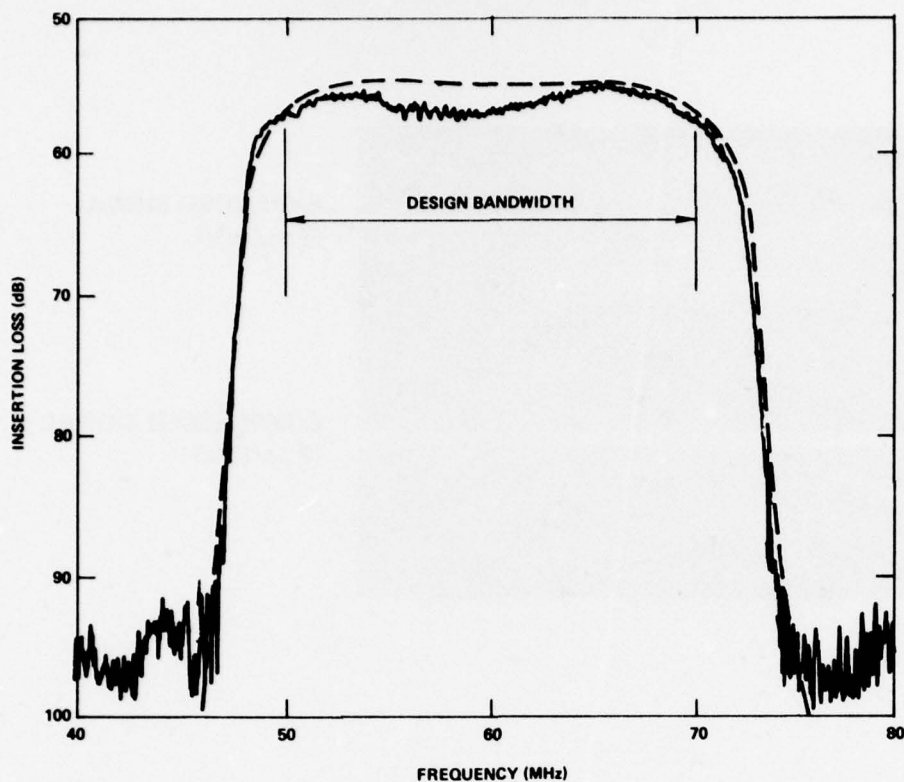
FIG. 1. Measured (solid line) and computed (dotted line) passband response of an RDA linear FM filter with F_c = 60 MHz, ΔF = 28 MHz and ΔT = 14 µs.
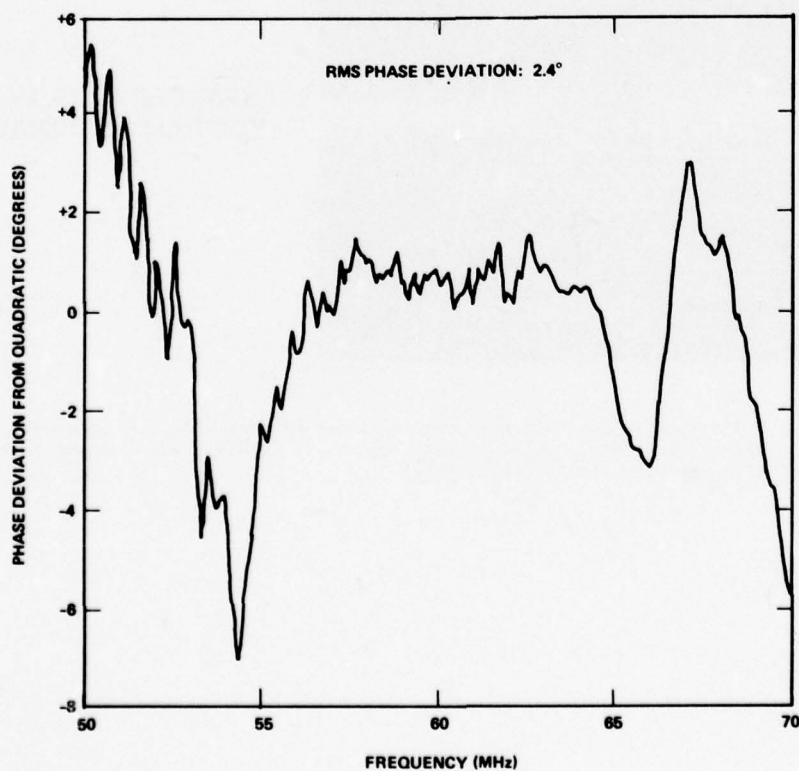
FIG. 2. Deviations of the measured phase from a quadratic fit vs frequency over the design bandwidth of 20 MHz. The rms phase deviation is 2.4°.

(a) EXPANDED SIGNAL (2 µs/DIV)

(b) COMPRESSED SIGNAL (2 µs/DIV)

(c) EXPANDED DISPLAY OF COMPRESSED SIGNAL (0.1 µs/DIV)

FIG. 3. Pulse compression test with two identical RDA devices. (a) expanded signal (2 µs/div), (b) compressed signal (2 µs/div), (c) expanded display of compressed signal (0.1 µs/div).

DISCUSSION

E.Stern

Could you comment on the relative merits of RDA filters as compared with conventional ID filters? In addition, could you comment on the relative cost?

Author's Reply

The cost of making the RDA device is very comparable to ID transducers. They are both single-step fabrication processes. The defect tolerance for RDA is less critical than for I.D. transducers.

In terms of performances, the ID filters will in general have a lower insertion loss due to the absence of the reflection loss by an amount in the order of 10 db, but the quality of the filter response will be much worse especially for high TB devices due to second-order effects such as bulk mode interference, regeneration, interelectrode reflections and a lower defect tolerance. Also, the ID filters require twice the substrate length for a given differential time delay. In short, all the advantages of the reflective groove devices apply to RDA devices as well.

P.Tournois

Can you comment on the fact that your results are not so good as have been obtained with groove technology?

Author's Reply

Results are not yet comparable with those obtained with reflective groove as we are in an earlier stage of development. In particular, the understanding of the velocity reduction in the array has to be improved and we are making progress in this way. Phase deviations have been 75 degrees, then 25 degrees, then 5 degrees, which makes an improvement.

P.Carr

In defence of I.D. transducers you set lower insertion loss because you have no reflection loss and also you have the temperature compensation of ST cut quartz.

Author's Reply

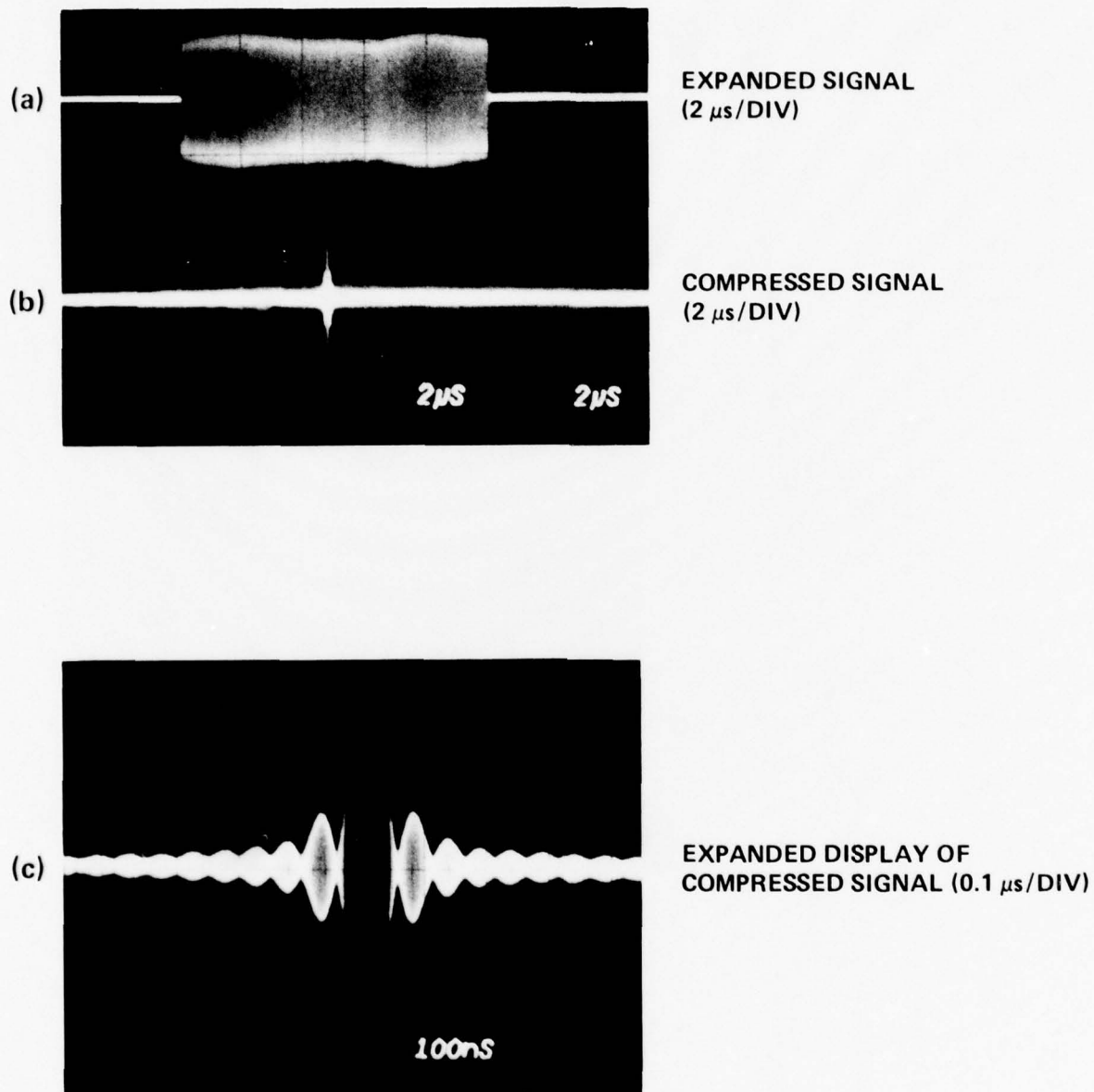The size of multireflection loss with the bandpass filter and with the linear FM filter has been around 10 to 11 db per reflection when we had the proper compensation. With thicker film we have observed the reflection loss will be 6 db per array with the proper compensation. The propagation delay in the principal direction, X direction, is temperature-compensated.

Maerfeld

You seem to have many problems in the characteristics of propagation of surface wave in your array. Why don't you use thinner metallic dots?

Author's Reply

The reason is that we want to minimize the insertion loss. If we have thin metallic dots the dispersion is less, the amount of compensation is less, but the insertion loss is higher.

I do not expect that the dispersion is more significant than in reflective groove devices.

M.Butler

The dispersive characteristics that you have shown have between 50 and 60 db insertion loss and you have not included compensation for metal thickness. Can you foresee an improvement for such a device up to 30 or 35 db? Would you not expect it to be very sensitive to metal thickness of the material? If so, could not the control of metal thickness be too demanding?

Author's Reply

The insertion loss for the transducers which were untuned account for 37 db. The reflection loss added gives a 55 db total loss, slightly under 10 db per array. With tuned transducers the insertion loss could be reasonable.

As regards control of metal thickness, we feel that it can be obtainable.

INFLUENCE DE L'ACCÉLÉRATION SUR DES OSCILLATEURS A ONDES ACOUSTIQUES DE SURFACE

par Michel VALDOIS*, Patrick LEVESQUE*, Pierre HARTEMANN**

*Office National d'Etudes et de Recherches Aérospatiales (ONERA)
92320 Châtillon - France

**Thomson - CSF, Laboratoire Central de Recherches, 91401 Orsay - France

L'influence de l'accélération sur des oscillateurs utilisant une ligne à retard ou un résonateur à ondes élastiques de surface a été étudiée dans une plage s'étendant de 0 à 30 g. Les oscillateurs à ondes de surface ont été disposés sur un bras rotatif de vitesse variable.

Les substrats piézo-électriques utilisés sont en quartz de coupe ST et Y. Plusieurs directions d'accélération ont été expérimentées. La déviation de fréquence due à l'accélération est en première approximation une fonction linéaire de l'amplitude de l'accélération. Par exemple une déviation de fréquence d'environ 10 Hz/g (g = 9,81 m/s²) à 105 MHz a été mesurée pour un oscillateur à ligne à retard en quartz de coupe ST lorsque la direction d'application de l'accélération est parallèle à celle de propagation des ondes de surface. De même, une direction de plus faible sensibilité a été observée, pour laquelle la sensibilité est inférieure à 1 Hz/g.

Des résultats similaires ont été trouvés pour un oscillateur à ligne à retard en quartz de coupe Y. Avec un résonateur à ondes de surface constitué de deux transducteurs déposés dans une cavité définie par deux réseaux périodiques réfléchissants, la sensibilité maximale est de 48 Hz/g à 124 MHz pour une direction d'accélération approximativement perpendiculaire à la surface de propagation.

Ces résultats montrent que les oscillateurs à ondes élastiques de surface, sous leur forme actuelle, sont beaucoup plus sensibles à l'accélération que les oscillateurs à ondes de volume. Le rapport d'augmentation de la sensibilité relative est de l'ordre de 100.

♦

INFLUENCE OF ACCELERATION ON SURFACE ACOUSTIC WAVE OSCILLATORS

The influence of acceleration on oscillators using a delay line or a resonator with Surface Acoustic Waves has been studied in the 0 - 30 g range. The surface wave oscillators were installed on a variable speed rotating arm.

The piezo-electric substrates used are in quartz of ST and Y cut. Several directions of acceleration were experimented. The frequency shift due to acceleration is, as a first approximation, a linear function of the acceleration magnitude. For example, a frequency shift of about 10 Hz/g (g = 9.81 m/s²) at 105 MHz has been measured on a delay line oscillator of ST cut, when the acceleration direction is parallel to that of surface wave propagation. In the same manner, a direction of lower sensitivity has been observed, for which sensitivity is lower than 1 Hz/g

Similar results have been found on a delay line oscillator in quartz of Y cut.

With a surface wave resonator made of two transducers placed in a cavity defined by two periodic reflecting networks, the maximum sensitivity is 48 Hz/g at 124 MHz for an acceleration direction perpendicular to the propagation surface.

These results show that the acoustic surface wave oscillators, in their present form, are much more sensitive to acceleration than the volume wave oscillators. The ratio of relative sensitivity increase is of the order of 100.

1 - INTRODUCTION

L'intérêt pour les oscillateurs à ondes élastiques de surface a été très grand depuis quelques années, car ces oscillateurs fonctionnent à très hautes fréquences fondamentales.

Cependant leurs caractéristiques ne sont pas encore complètement définies. Pour certaines applications aérospatiales, il est important de connaître leur comportement en présence d'accélération statique. Certains résultats expérimentaux récents, concernant l'influence de l'accélération sur les oscillateurs à ligne à retard ou à résonateur à ondes élastiques de surface, sont publiés dans cet article. Après une description des méthodes de mesure utilisées, des résultats d'expériences portant sur des oscillateurs à ligne à retard en quartz de coupe ST et Y, et des oscillateurs à résonateur en quartz de coupe ST sont donnés.

2 - METHODES DE MESURE

Le dispositif à ondes de surface et son amplificateur de rebouclage sont enfermés dans une enceinte à vide (10^{-6} mbar), thermiquement isolée dans le cas de quartz de coupe ST, pour obtenir une meilleure stabilité de température. En effet, la température du point de renversement de la caractéristique fréquence-température est, pour la coupe étudiée, située aux environs de 20° C où la thermorégulation est très difficile. Lorsque cette température est comprise entre 40° C et 100° C, les dispositifs acoustiques sont disposés dans un thermostat qui stabilise la température avec une précision de 0,1° C. Les oscillateurs sont fixés sur un bras tournant montré sur la figure 1, à 1,5 m de l'axe de rotation. L'amplitude de l'accélération est ajustée par variation de la vitesse de rotation. Un contacteur tournant VHF est utilisé pour transmettre le signal de l'oscillateur aux appareils de mesure.
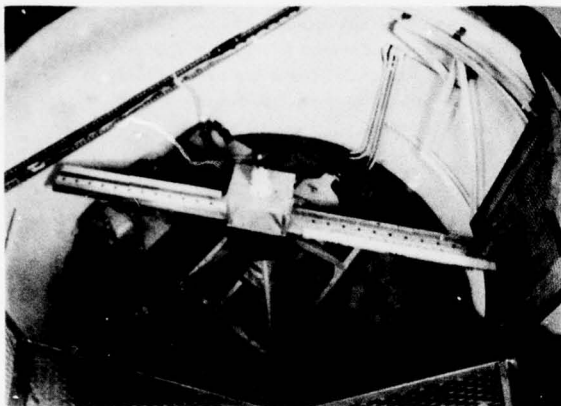

Fig. 1 — Bras rotatif utilisé pour appliquer des accélérations centrifuges. La longueur utile du bras est de 1,5 m.

Deux méthodes de mesure ont été employées. La première consiste à mesurer directement les variations de fréquence qui sont mises en mémoire dans un calculateur. La précision absolue de mesure dans ce cas est de 1 Hz pour un temps de comptage d'une seconde. Cette précision a été améliorée par une deuxième méthode qui consiste à asservir la fréquence d'un synthétiseur sur la fréquence du signal à mesurer. Le schéma de principe est donné figure 2.
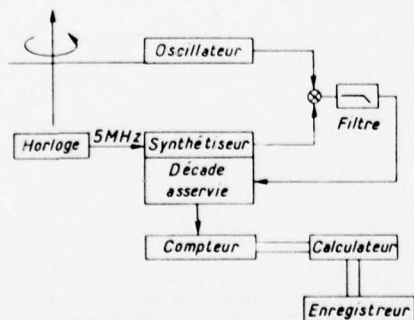

Fig. 2 — Schéma synoptique du montage de mesure des déviations de fréquence par asservissement.

La fréquence de l'oscillateur à ondes de surface est comparée à une référence issue d'un synthétiseur. Le battement basse fréquence obtenu est converti en une tension qui pilote le synthétiseur dans une plage de fréquence donnée. La précision absolue de cette méthode est de 0,1 Hz pour un temps de comptage d'une seconde.

3 - RESULTATS EXPERIMENTAUX

Les comportements de deux oscillateurs à ligne à retard utilisant des substrats en quartz de coupe ST et Y ont été d'abord étudiés (figure 3), puis celui d'un oscillateur à résonateur en quartz coupe ST. Dans tous les cas, les substrats ont été collés avec une résine époxy sur le fond d'un boitier métallique, pour pouvoir supporter de fortes accélérations.
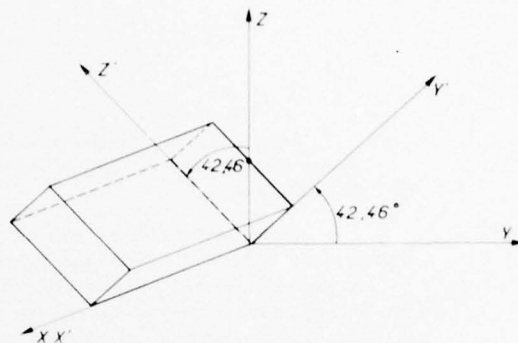
Fig. 3 — Orientation cristallographique de la coupe ST.

3.1. - Ligne à retard en quartz coupe ST

La ligne est constituée de trois transducteurs en peignes (figure 4). Deux de ces transducteurs sont échantillonnés (HARTEMANN P, 1971), et servent à l'oscillation, le troisième permettant de prélever le signal. Les deux transducteurs échantillonnés se composent de transducteurs élémentaires comportant trois dents de 7,5 μm de large. La distance entre 2 éléments adjacents est égale à 5 longueurs d'ondes acoustiques pour un de ces transducteurs et à 7 longueurs d'ondes pour l'autre. Le nombre d'éléments est de 90 pour le transducteur émetteur et 65 pour le récepteur. Les pertes d'insertion entre ces deux transducteurs sont de 12 dB à 105,2 MHz avec une bande passante à 3 dB de 100 kHz. Le retard est de 4,67 μs . L'amplificateur dont l'utilisation est nécessaire pour réaliser un oscillateur (LEWIS, MF, 19) est connecté entre ces deux transducteurs, le transducteur central étant l'émetteur. Le troisième transducteur, de structure conventionnelle, est utilisé pour coupler la boucle d'oscillation au circuit extérieur. Il est constitué de 150 dents, et la perte d'insertion entre le transducteur central et ce troisième transducteur est de 11,5 dB. Cet oscillateur présente de bonnes caractéristiques spectrales de bruit de phase comme le montre la figure 5. A 10 kHz de la fréquence nominale, le niveau du palier de bruit de phase est de - 155 dB/Hz.

Les accélérations ont été appliquées dans trois plans orthogonaux associés au substrat : X'OZ', Y'OZ', X'OY', X' étant la direction de propagation parallèle à l'axe cristallographique X, et Y' l'axe orthogonal du substrat.

Pour une accélération d'amplitude constante de 5 g, les déviations maximales de fréquence enregistrées dans les plans X'OZ', Y'OZ' et X'OY' sont respectivement de 60 Hz (12 Hz/g), 80 Hz (16 Hz/g) 50 Hz (10 Hz/g) soit en valeurs relatives $1,14.10^{-7}/g$ $1,52.10^{-7}/g$, $9,50.10^{-8}/g$.

Pour une direction donnée quelconque, la variation de fréquence est sensiblement proportionnelle à l'amplitude de l'accélération appliquée. A titre d'exemple, la figure 6 représente ces variations de fréquence observées pour la direction X' et dans une plage de 25 g pour laquelle la sensibilité relative moyenne est de l'ordre de $8,75.10^{-8}/g$.

Fig. 6 — Variation de fréquence en fonction de l'amplitude de l'accélération d'un oscillateur à ligne à retard à ondes acoustiques de surface à 105 MHz.
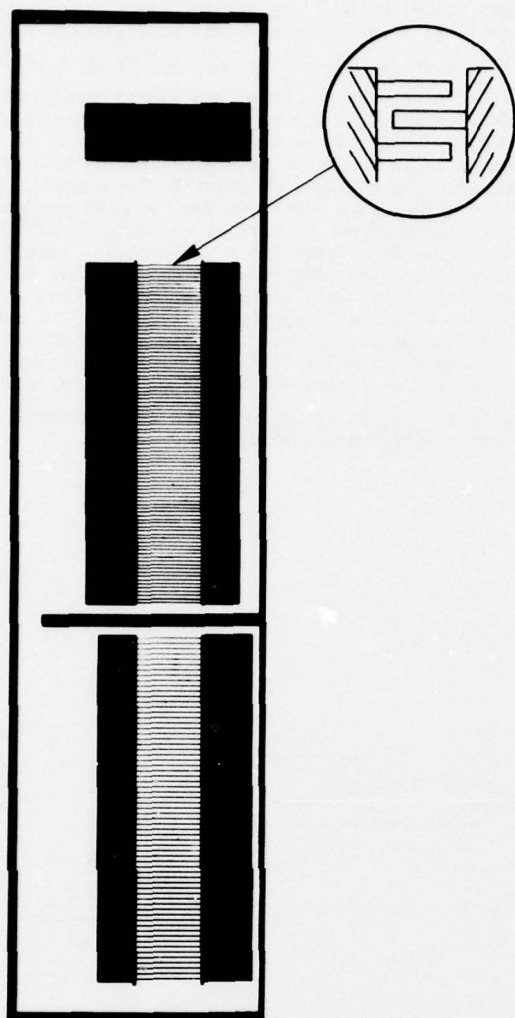
Fig. 4 — Vue de la ligne à retard à 105 MHz. La longueur de la plaquette de quartz est de 40 mm.
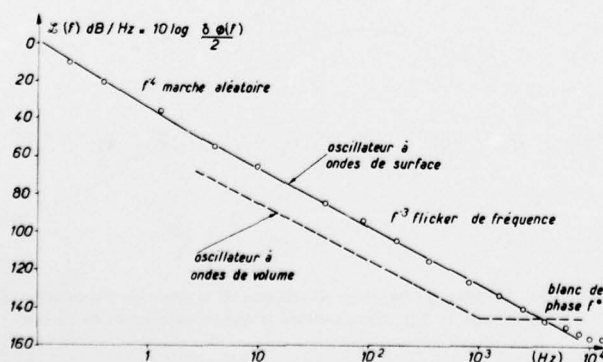
Fig. 5 — Densité spectrale unilatérale du bruit de phase d'un oscillateur à ligne à retard à ondes acoustiques de surface à 105 MHz.

3.2. - Ligne à retard de coupe Y

Une plaquette de quartz coupe Y a été utilisée avec une direction de propagation voisine de X + 35°. Dans ce cas, le point de renversement de la caractéristique fréquence-température est de 55° C avec un coefficient de $17.10^{-9}/(°C)^2$ (HARTEMANN, P, 1976). Cette coupe a donc été utilisée pour sa stabilité en température meilleure que celle de la coupe ST. En employant la même configuration de transducteur, la fréquence centrale est de 111,5 MHz et la perte d'insertion entre les deux transducteurs échantillonnés est de 17,5 dB.

Soumise à l'accélération dans trois plans orthogonaux X'OZ', Y'OZ' et X'OY', la ligne a des sensibilités maximales mesurées par retournement respectivement égales à 12 Hz/g, 17 Hz/g, 17 Hz/g. Les déviations relatives maximales sont comprises pour ces trois plans entre $1,07 \cdot 10^{-7}/g$ et $1,52.10^{-7}/g$.

3.3. - Résonateur coupe ST

Un oscillateur à résonateur à ondes élastiques de surface a été soumis à des accélérations. Il est constitué de deux transducteurs déposés entre deux réseaux périodiques comportant 630 traits implantés

ioniquement dans du quartz coupe ST (HARTEMANN, P, 1975), (figure 7). Les pertes d'insertion à la fréquence centrale de 124,4 MHz sont de 4 dB. Le point de renversement de la caractéristique fréquence-température est de 45° C, car cette température est augmentée par suite des effets de l'implantation d'ions.

On constate que la courbe est affectée par une non-linéarité pour des accélérations croissantes. La sensibilité relative pour une accélération inférieure à 10 g est de l'ordre de $3,1.10^{-7}/g$ et décroît à $2,3.10^{-7}/g$ de 30 à 40 g. Des essais complémentaires de retournements ont permis d'enregistrer continuement la variation de fréquence en fonction de l'angle que fait la direction de l'accélération avec une direction quelconque du cristal. Par exemple, l'accélération a été appliquée dans un plan perpendiculaire au substrat et faisant un angle de 30° avec la direction de propagation X' (figure 9). Dans ce plan, la variation de fréquence est une fonction sinusoïdale de l'angle que fait l'accélération avec la normale au substrat, comme le montre la figure 10 ; il existe donc une direction pour laquelle la variation de fréquence est nulle (Valdois, M, 1977). La valeur maximale de la sensibilité est de 48 Hz/g, soit $3,86.10^{-7}/g$, et est très proche de la direction de sensibilité maximum maximorum.

Fig. 7 — Vue du résonateur à ondes élastiques de surface.

La sensibilité accélérométrique a été étudiée dans une direction donnée, pour différentes amplitudes d'accélération jusqu'à 40 g. Cette direction fait un angle de 15° avec l'axe Y' et la variation de fréquence résultante est représentée figure 8.
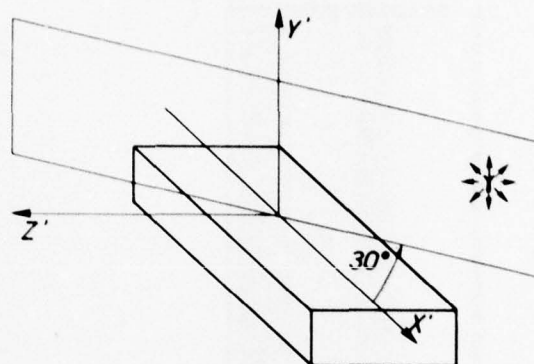
Fig. 9 — Représentation du plan étudié en accélération pour le résonateur à ondes de surface.
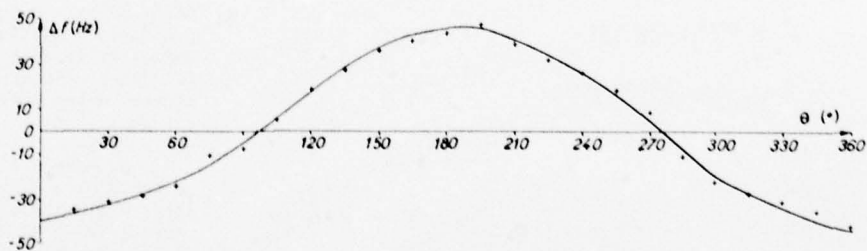
Fig. 10 — Variation de fréquence en fonction de la direction d'application d'une accélération de 1 g d'un oscillateur à résonateur à ondes de surface.
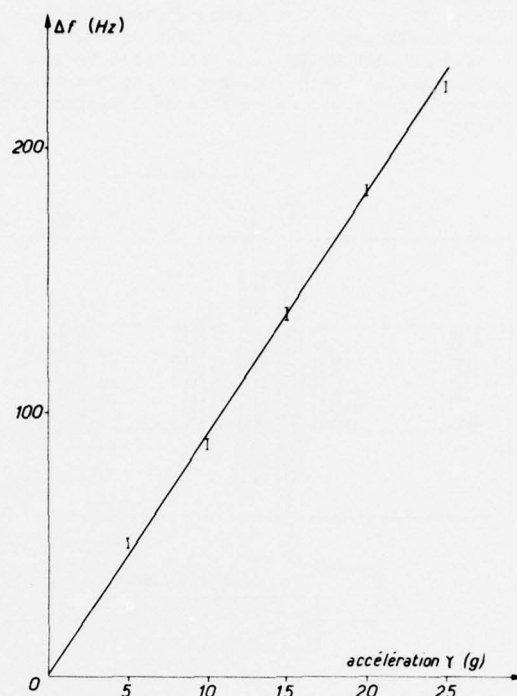

Fig. 8 — Variation de fréquence en fonction de l'amplitude de l'accélération d'un oscillateur à résonateur à ondes acoustiques de surface à 124 MHz.

4 - CONCLUSION

Ces premiers résultats expérimentaux montrent
que l'influence des accélérations sur les oscillateurs
à ondes élastiques de surface peut être très impor-
tante pour certaines directions. Ainsi, des sensi-
bilités relatives de $1,52.10^{-7}/g$ pour un oscillateur
à ligne à retard et de $3,86 \ 10^{-7}/g$ pour un oscillateur
à résonateur ont été observées. Les résonateurs sem-
blent plus sensibles que les lignes à retard. Cepen-
dant, une exploration complète des directions d'ap-
plication de l'accélération doit être effectuée pour
pouvoir comparer valablement les sensibilités des
lignes à retard et des résonateurs.

Les déviations de fréquence sont des fonctions
approximativement linéaires de l'amplitude de l'ac-
célération.

Nous avons montré aussi qu'il existait des
directions d'influences nulles dont les caractéris-
tiques sont actuellement étudiées.

Comparés aux oscillateurs à ondes élastiques
de volume, les oscillateurs à ondes de surface sont
environ 100 fois plus sensibles aux accélérations.

REMERCIEMENTS

REFERENCES

HARTEMANN, P, 1977
"Narrow bandwidth Rayleigh wave filters". Elec-
tronics letters, 7, 674.

HARTEMANN, P, 1975,
"Acoustic surface wave resonator using ion
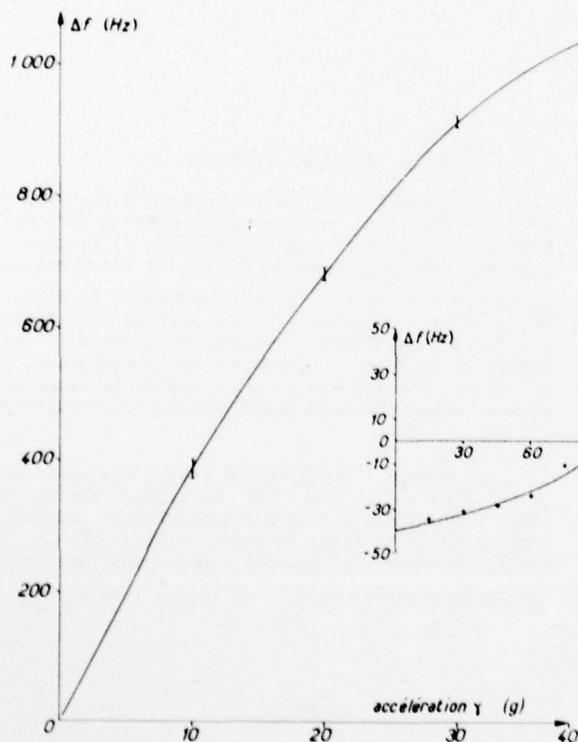implanted gratings".
Ultrasonics Symposium Proceedings, 303.

HARTEMANN, P, 1976,
"Oscillator temperature stability change induced
by ion implantation"
Ultrasonics Symposium Proceedings, 240.

LEWIS, MF, 1973,
"Some aspects of SAW oscillators"
Ultrasonics Symposium Proceedings, 344.

VALDOIS, M, GAGNEPAIN, JJ, BESSON, J, 1974,
"Influence of environment conditions on a
quartz resonator".
Proceedings of the 28th Annual Symposium on
Frequency Control, 19.

VALDOIS, M, JANIAUD, D,
"Existence d'un plan d'influence accélérométri-
que nulle pour les résonateurs à quartz".
Mesures - To be published.

DISCUSSION

P.Carr

Have you attempted to decrease the vibration sensitivity of your device by increasing the thickness of your slide or mounting technique? Have you tried to exploit this feature of the SAW?

Réponse d'Auteur

J'ai parlé de sensibilité aux accélérations statiques et non aux vibrations.

En ce qui concerne la seconde question nous jouons sur la configuration du système pour rechercher les meilleures dispositions.

P.Tournois

En vibrations Hughes Aircraft a publié des résultats où la sensibilité des ondes de surface (résonateur en particulier) aux vibrations est inférieure à celle des ondes de volume. Vous montrez maintenant que la sensibilité aux accélérations est supérieure à celle des ondes de volume. A quelle fréquence doit-on trouver l'égalité?

Réponse d'Auteur

On ne connait pas encore d'écrans accélérations statiques, c'est la raison pour laquelle nous avons commencé par elles. Nous continuerons par les vibrations et nous vous communiquerons les résultats.

Maerfeld

Avez vous une explication physique à cette sensibilité observée?

Réponse d'Auteur

Nous n'avons pas encore d'explications physiques.

Nous attendons des résultats d'études théoriques en cours à BESANCON et à l'ONERA pour voir si elles correspondent aux observations.

Toutefois il y a un point très délicat c'est la transformation des accélérations en force par l'intermédiaire de l'ancrage des substrats sur leurs supports.

E.Stern

Was the crystal attached to substrate at every point of interface with epoxy?

Could this contribute to frequency sensitivity, due to differences in elasticity of crystal and substrate?

Réponse d'Auteur

Le cristal est collé à l'aide d'une résine époxy sur toute sa surface de base. Une pression est exercée au moment du collage afin de répartir uniformément la colle, mais aucun contrôle de cette répartition n'a été effectué.

Les conditions de fixation du cristal de quartz sur le support définissent des contraintes en tout point de celui-ci, qui varient pour un même quartz en fonction des paramètres suivants:

— forme et nature du support

— interface cristal/support.

En effet, seules les variations des contraintes peuvent être retenues pour expliquer la sensibilité de fréquence compte tenu que, pour un cristal de 3 mm d'épaisseur, les composantes des déplacements sont négligeables au niveau de l'interface cristal/support. En conséquence, toute variation des conditions de fixation du cristal sur le support entraine une modification des contraintes et de leur répartition, et des coefficients élastiques déterminant la vitesse de propagation de l'onde. D'autres types de supports sont actuellement à l'étude afin de modéliser l'influence de tels paramètres sur les sensibilités accélérométriques.

TUNABLE MAGNETOELASTIC SURFACE WAVE OSCILLATORS

G. VOLLUET

THOMSON-CSF Research Center

Domaine de Corbeville - 91401 Orsay (France)

SUMMARY

A new technique for achieving the frequency control of S.A.W. oscillators with an external magnetic field has been experimented. Magnetoelastic Rayleigh wave delay lines using ZnO transducers have been fabricated with Gd-Ga Y.I.G. epitaxial films grown on G.G.G. substrates. An important phase velocity shift has been obtained as a function of an external biasing magnetic field through a magnetoelastic interaction. Feedback oscillators have been implemented using this kind of delay lines. Then a frequency change up to 1650 ppm around 211.5 MHz and 1800 ppm around 339.5 MHz has been obtained by sweeping a biasing magnetic field over a 70 Oe range.

1. INTRODUCTION

Over the past few years, there has been a growing interest in surface acoustic wave (S.A.W.) oscillators. For, the fundamental frequency of bulk wave resonators obtained by conventional technique is limited to about 40 MHz, whereas S.A.W. delay lines can operate over a greater frequency range (up to 2 GHz). Generally a S.A.W. oscillator consists of a loop containing a delay line and an amplifier. An important feature of oscillators is their frequency modulation capability. But with conventional S.A.W. delay line the delay and the frequency response of the line are constants and cannot be continuously changed. However the frequency modulation has been achieved by inserting a phase shifter in the loop (LEWIS, M.F., 1974). With such a technique, the oscillator operates at frequencies more or less far from the centre of the frequency response of the delay line. From this point of view it would be better to modify the velocity along the acoustic path, then the delay and the frequency response would vary together. This is achieved using the new method described in this paper.

Recently a magnetic control of the acoustic velocity on $LiNbO_3$ has been obtained by depositing a magnetostrictive thin nickel film (GANGULY, A.K. et al., 1975). A relative phase velocity shift of 2.10^{-4} at 210 MHz was measured in this case. It has been shown theoretically (PAREKH, J.P. et al., 1976) that new "giant magnetostrictive" rare-earth iron compounds (such as $Tb_{.23} Dy_{.73} Fe_2$) could give stronger dispersive effects. However, the acoustic attenuation is drastically increased with such conductive materials. One can think to overcome this limitation by the use of non-conductive magnetic garnets. A first experiment was conducted with bulk garnet (LUNDSTROM, M.S. et al., 1974). Unfortunately the internal magnetic field in this geometry is not uniform. This effect limits the interaction zone, then the relative phase shift is weak.

A quite different structure to keep out of this trouble is reported in this paper. The acoustic propagating medium is an epitaxial magnetic garnet. When a magnetic field is applied in the plane of the film, the absence of demagnetizing effects leads to strong interaction all over the acoustic path, then a maximum phase shift can be obtained.

The first part of this paper describes the structure of the magnetoelastic delay line. In the second part results about phase delay variations are given and the best geometry to obtain strong dispersion is discussed. In the final part prelimenary results about oscillator frequency shift versus magnetic field are reported.

2. EXPERIMENTAL RESULTS AND DISCUSSION

2.1. Magnetoelastic delay line

A schematic of the experimental arrangement of the magnetoelastic delay line is shown in figure 1. The acoustic propagation medium is a (1,1,1) doped Yttrium-Iron-Garnet (Y.I.G.) film grown on a non-magnetic substrate of gallium-gadolinium-garnet (G.G.G.) by the liquid phase epitaxy process. Gallium and gadolinium are substituted in Y.I.G. to obtain the so-called <<easy-plane-film>> (the composition in this case is $Gd_{.45} Y_{2.55} Ga_{.9} Fe_{4.1} O_{12}$).This material has a cubic symmetry. The stress induced by the epitaxy (a compression in this case) forces the magnetization to lie in the film plane by the way of magnetostriction. Then, the film can be magnetically saturated in its plane by a weak external field. A sputtered piezoelectric zinc oxide layer and conventional interdigital transducers are used to launch and receive acoustic surface waves. Then magnetoelastic interactions were controlled by an external biasing magnetic field $\vec{H}$.

Two devices working at different frequencies have been tested. The first one (device No1) has been fabricated with a 25 μm thick Gd-Ga Y.I.G. (1,1,1) film. A 2 μm thick ZnO layer has been removed between the interdigital transducers to decrease acoustic losses. The acoustic propagation was chosen along the <1,1,$\bar{2}$> direction and the frequency has been close to 211 MHz. Then the insertion loss have been 18 dB with tuned transducers at magnetic saturation and the delay has been 3.7 μs. The Gd - Ga Y.I.G. (1,1,1) film thickness has been 9 μm and the surface wave propagation has been along <1,$\bar{1}$,0> in the device No2. With 1 μm thick ZnO layer and an acoustic wavelength of 10 μm, the centre frequency and the insertion loss have been 339.9 MHz and 22 dB respectively for a 1.86 μs delay. In order to obtain a better efficiency with the configuration of device No2 , the ZnO thickness would be 1.5 μm according to the theoretical prediction.

An external magnetic field has been applied in the plane of the film. The direction and the intensity of the biasing field controls the magnetoelastic interactions between the Rayleigh wave and the magnetic film.

2.2. Phase shift measurement

The variations of the phase delay as a function of the biasing magnetic field have been measured. These variations have been determined, the field being slowly decreasing. The phase velocity shift of the device No1 for different field directions is shown in figure 2. The maximum of the shift can only be obtained with a well defined magnetic field direction which is a <1,1,0>- type direction, the nearest of the acoustic propagation $\vec{k}$. There are four equivalent field directions in device No1 which are $\pm$ 30° and + 150° off the $\vec{k}$ direction. Results are similar for each one of these directions. The maximum phase shift is rapidly decreasing when the field direction varies from a <1,1,0> - type direction as shown in figure 2. For the device No1, with the best configuration, the measured relative phase velocity shift ($\Delta V/V$) is about 3.10^{-3} when the magnetic field is swept over 55 Oe - 120 Oe.

Associated to this phase shift there is an increase in acoustic attenuation. Theses variations are given in figure 3 for device No1 (solid line curve); the reference level is the acoustic signal at high magnetic field value. An extra-attenuation for the surface acoustic wave (more than 30 dB) has been measured between 25 Oe and 55 Oe. Such an attenuation has been explained by some leaky character of the acoustic wave according to calculations performed with a simpler configuration (PAREKH, J.P., et al., 1974). The relative phase velocity shift has been plotted in the same figure (dashed curve).

Experiment conducted with device No2 has shown greater total phase velocity shift : about $5.5 \, 10^{-3}$ for a 40 Oe to 120 Oe field variation as shown in figure 4 (dashed line curve). In this case the best magnetic field direction is along the acoustic propagation : $<1,\bar{1},0>$. A similar increase of the acoustic attenuation has also been observed.

With such a symmetry ((1,1,1) film plane) and taking into account the magnetocristalline anisotropy, exact analysis leads to untractable calculations. However to get an insight by a simple way into the conditions for obtaining the stronger interaction, an effective magnetic field has been associated with the Rayleigh wave through the magnetoelastic energy density (VOLLUET, G., et al., 1976). Moreover gyromagnetic resonances can be measured and calculated as a function of the biasing field for different directions. Then two conditions need to be satisfied to get a strong interaction. First a gyromagnetic resonance must exist at the expected frequency for the considered biasing field direction. Second, the effective magnetic field must be important to efficiently excite this resonance. The experimental results are in good agreement with the above hypothesis.

2.3. Oscillator characteristics

These delay lines have been used to implement S.A.W. oscillators as shown in figure 1. When the amplifier gain is greater than the insertion loss,the oscillator operates at a frequency (f) which satisfies the well known condition:

$$2 \pi f \frac{L}{V} + \phi_E = 2 n \pi \quad \text{(n integer)}$$

where L is the pathlength between transducers, V the acoustic phase velocity, ϕ_E the electrical shift associated to the amplifier and the matched transducers.

The measured phase velocity shift ($\Delta V/V$) with a biasing magnetic field can be used to shift the oscillator frequency by Δf such that :

$$\frac{\Delta f}{f} = \frac{\Delta V}{V}$$

Then it is possible to change the frequency and always to operate at the centre frequency of the transfert function. The frequency response of the delay line No2 for different magnetic field values is shown in figure 5. The magnetic control of the band-pass centre can lead to good working conditions for a tunable oscillator with a relatively long delay line.

The spectrum of the wave transmitted by the oscillator is shown in figure 6-a. When the magnetic field is varied, the spectral response is always clean but the frequency is shifted as shown in the second picture (figure 6-b). The field has been swept for obtaining a shift over 300 KHz with oscillator No2.

The large phase velocity variation cannot be entirely used due to the extra acoustic attenuation (figures 3 and 4). To partly overcome this effect, the amplification was greater than the insertion loss. Good operating conditions have been obtained with 8 dB excess of gain for oscillator No1 and 10 dB excess for oscillator No2.Then one can hope to have a frequency change of 1900 ppm and 2200 ppm respectively (figure 3-4). Measurements have been done starting from a high magnetic field value ($\sim$ 140 Oe) and slowly decreasing the field. Results are shown in Fig. 7 and 8 for oscillators No1 and No2 respectively The general behaviour is like the same for these two devices. The measured total frequency shift of the oscillators have been respectively about 350 KHz and 620 KHz. It corresponds to a relative value of 1650 ppm over a field range of 65 Oe - 125 Oe with oscillator No1 and 1800 ppm over a field range 47 Oe - 125 Oe with oscillator No2. These quantities are slightly lower than the predictions because it needs a weak excess of gain to operate correctly.

Magnetic materials are generally influenced by temperature variations, mainly by the change in magnetization and anisotropy constant. To characterize the behaviour of the oscillator with temperature, the frequency drift of device No2 has been measured over 10°C - 80°C. In this experiment, the oscillator was operating at zero magnetic field to avoid some secondary effects due to magnetic field shift with temperature. A linear variation of frequency of -53 ppm have been measured.

3.   CONCLUSION

The ability to control the phase velocity of surface acoustic wave delay line is very important for further development of tunable S.A.W. oscillators. New technique is described in this paper for magnetically controlling the phase velocity of Rayleigh wave. Two devices have been experimented at different frequencies. A frequency change up to 1800 ppm has been measured with only 70 Oe for the magnetic field variations.

In such devices, a large frequency shift can be obtained with a long delay because the relative frequency change is equal to that of the phase velocity, then the variations are independant of the acoustic path length. Moreover the oscillating mode frequency is always at the centre of the frequency response of the line and mode hopping can be avoided.

Although magnetoelastic properties of Gd - Ga Y.I.G. are relatively weak, an important phase velocity shift has been measured. A noticeable increase of this effect could be probably obtained with more magnetostrictive thin garnet films. For exemple bulk Tb doped Y.I.G. are known to have stronger magnetostriction coefficients at room temperature. Then epitaxial films of such a composition would be a more suitable propagating medium for increasing the phase velocity shift.

Aknowledgements :

REFERENCES

GANGULY, A.K., DAVIS, K.L., WEBB, D.C., VITTORIA, C. and FORESTER, D.W., 1975, " Magnetically tuned surface-acoustic-wave phase shifter", Elec. Letters, 11, P. 610-611.

LEWIS, M.F., May 1974, "Oscillators - the next succesful surface acoustic wave device ? ", Ultrasonics, P.115-123.

LUNDSTROM, M.S. and ROBBINS, W.P., 1974, Ultrasonics Symposium Proceedings, (IEEE cat. N° 74 CHO 896-1SU), pp. 348-351.

PAREKH, J.P. and BERTONI, H.L., January 1974, "Magnetoelastic Rayleigh waves propagating along a tangential bias field on a YIG substrate", J. Appl. Phys., 45, N° 1, pp.434-445.

PAREKH, J.P. and SAVAGE, H.T., 1976, "Rayleigh wave in strongly magnetostrictive materials", Ultrasonics Symposium Proceedings, (IEEE cat. N° 76 CH1 120- 5SU), pp. 673-675.

VOLLUET, G., DESORMIERE B., and AULD, B.A., 1976, "Magnetoelastic effects associated with elastic surface wave propagation in epitaxial garnet films", A.I.P. Conf. Proceedings, 34, pp. 274-276.
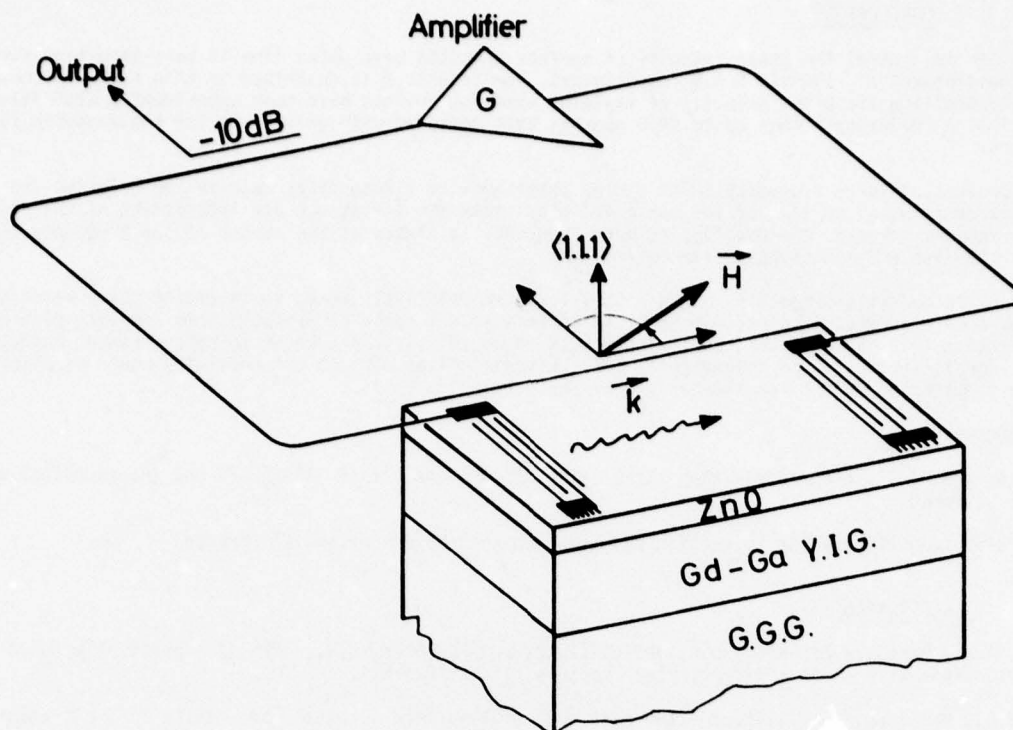
Figure 1 - Schematic of the magnetoelastic delay line oscillator
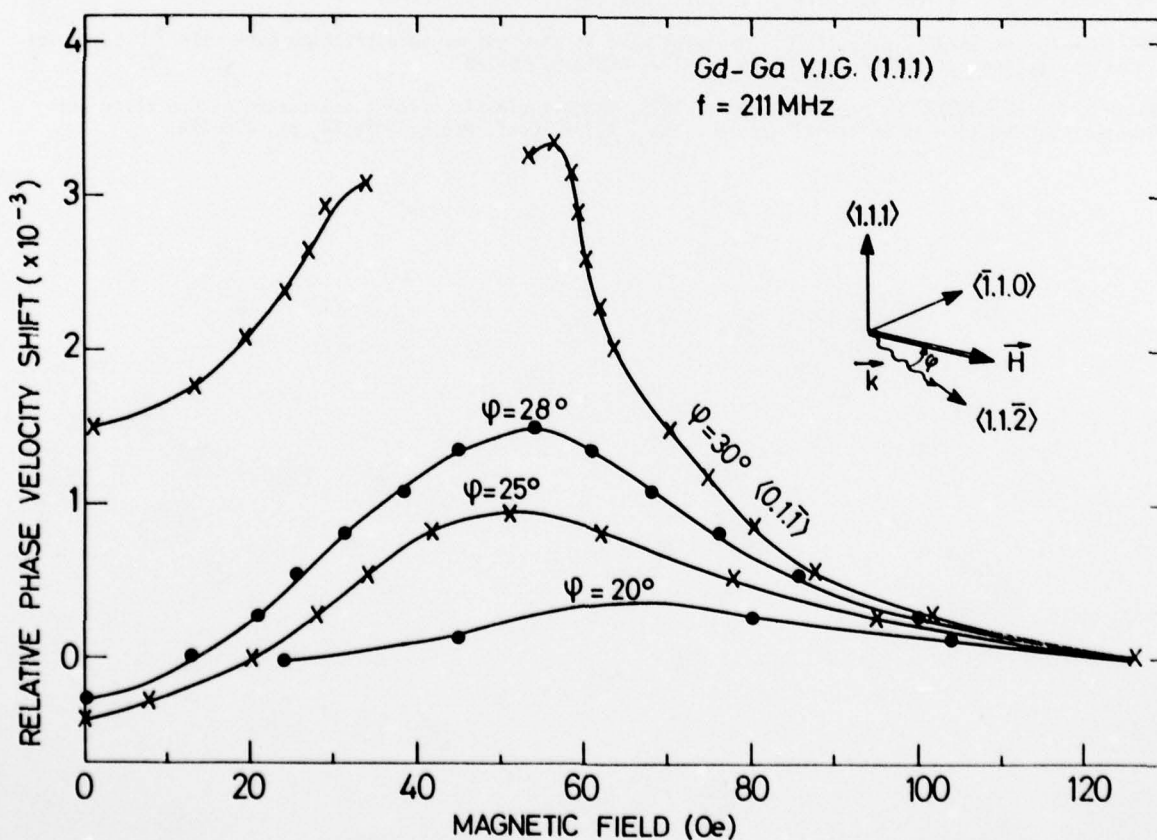
Figure 2 - Relative phase velocity shift as a function of the applied field for different field orientations (device No1)
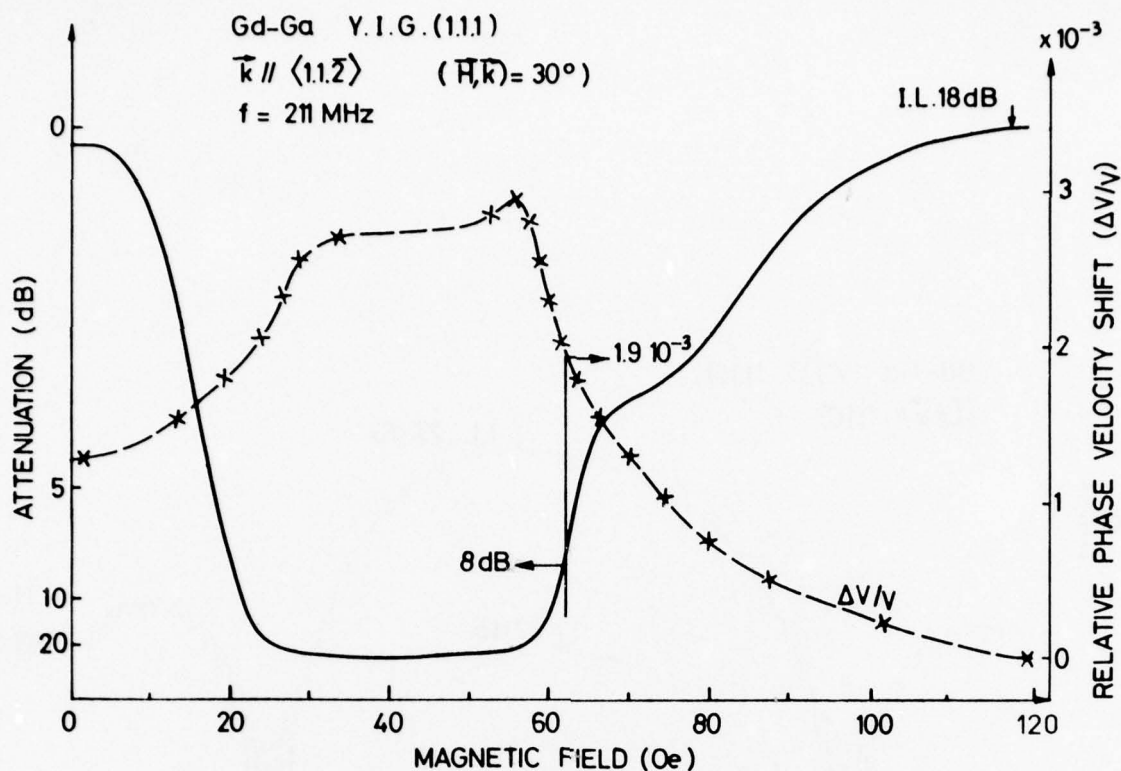
Figure 3 - Acoustic attenuation (solid line) and phase velocity shift (dashed line) as a function of an in-plane magnetic field - (device No1)
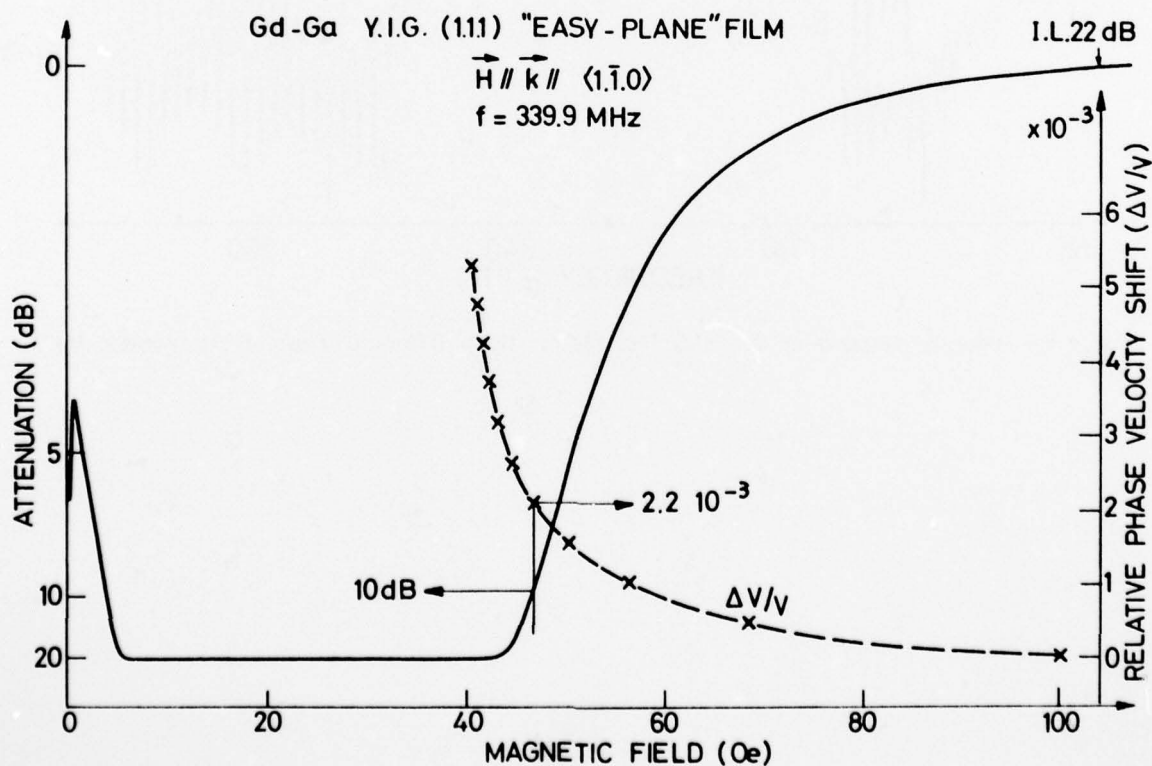
Figure 4 - Acoustic attenuation (solid line) and phase velocity shift (dashed line) as a function of an in plane magnetic field - (device No2)
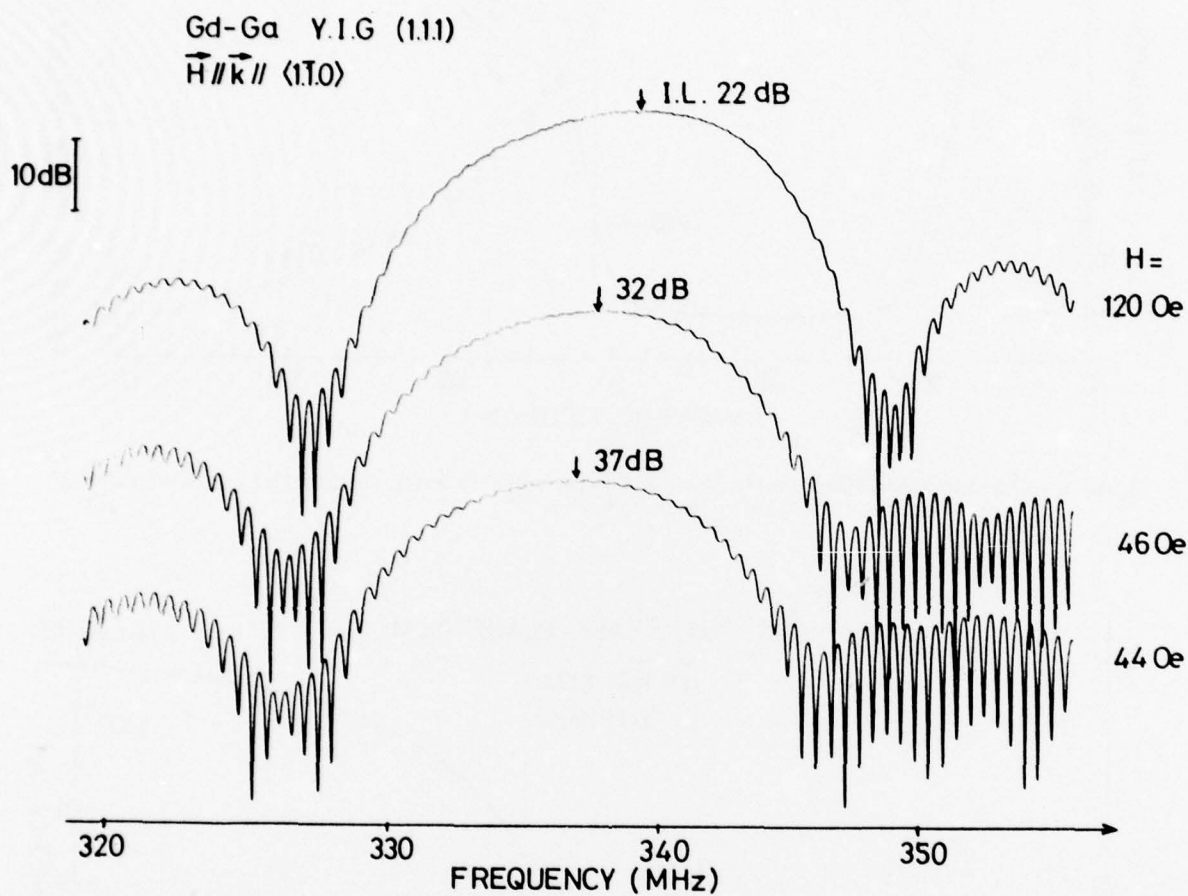
Figure 5 - Frequency response of the delay line No2 for three different values of the magnetic field.

OUTPUT SPECTRA

10 dB.↕
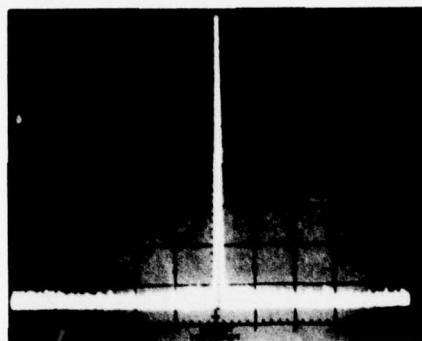
horizontal scale
200 kHz./div.

-a-

FREQUENCY SHIFT

10 dB.↕
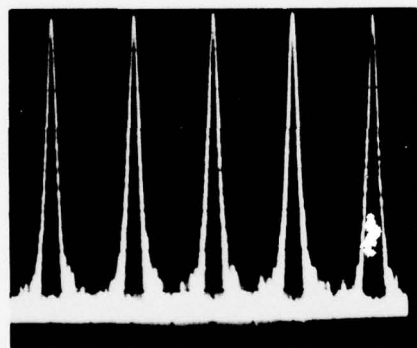
horizontal scale
50 kHz./div.

52 56 62 75 100

Magnetic field (Oe.)

-b-

Figure 6 - Output spectra of the oscillator No2
a - The magnetic field has a fixed value
b - Each spectrum corresponds to a well defined magnetic field value.
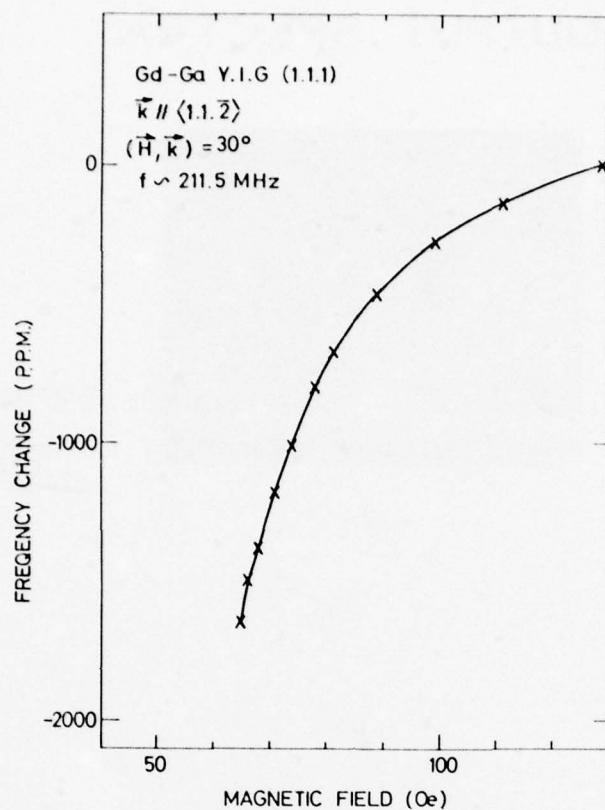
Figure 7 - Frequency shift versus in-plane magnetic field (oscillator No1)
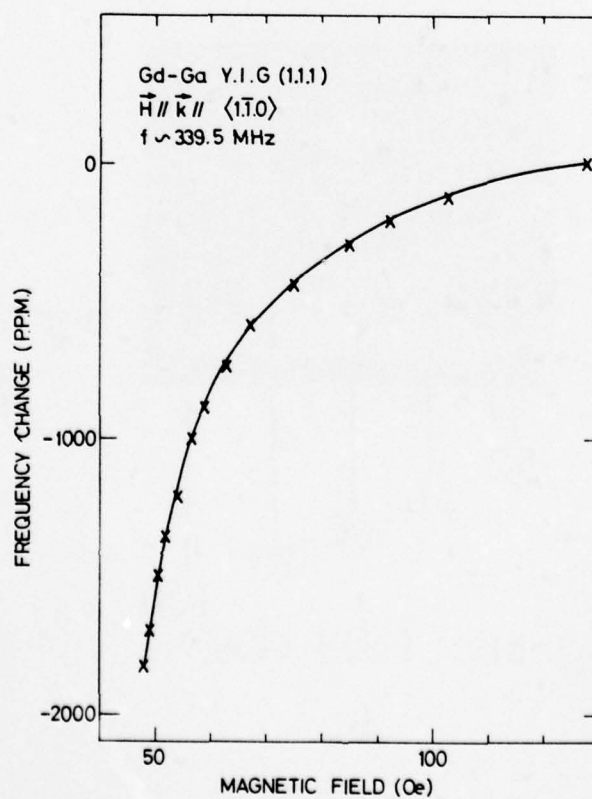
Figure 8 - Frequency shift versus in-plane magnetic field (oscillator No2)
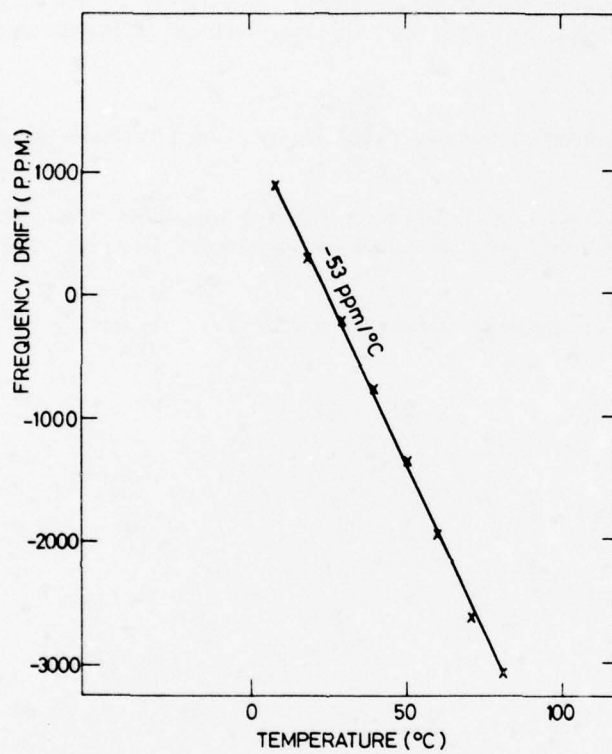
Figure 9 – Frequency drift of the oscillator No2 versus temperature.
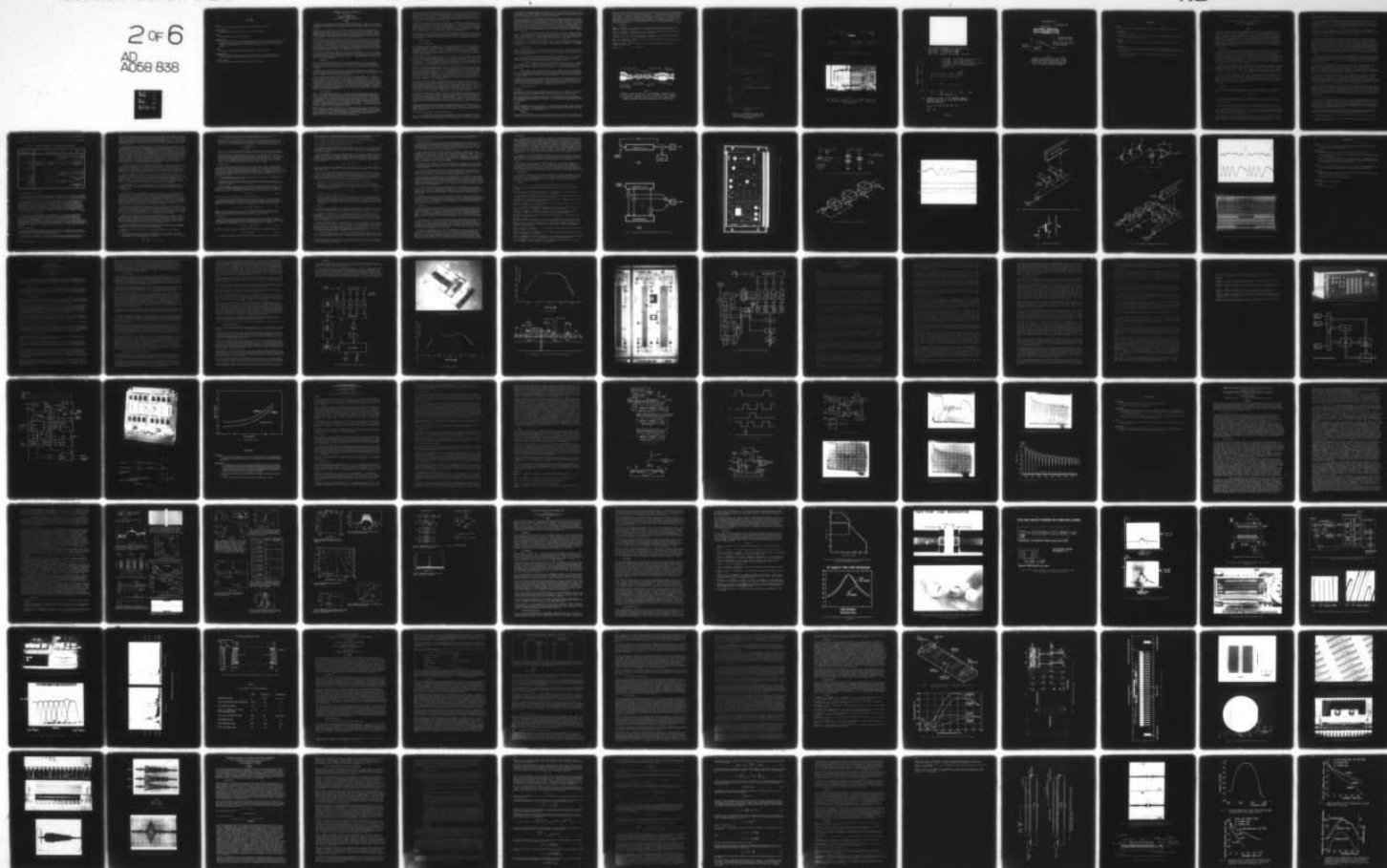
DISCUSSION

P.Carr

Have you observed any non-reciprocal surface accoustic wave due to the magnetic film?

Réponse d'Auteur

Il a été montré récemment que l'on a affaire à des phénomènes non réciproques.

Voles

How reproducible are you expecting this material to be?

Réponse d'Auteur

Ces études sont à leur début. Les inconvénients sont très importants: Fortes dérives en température, fortes pertes d'insertion + On envisage évertuellement des modifications des substrats. Il est possible que l'on obtienne des résultats plus satisfaisants dans un proche avenir. On peut obtenir une variation de vitesse à l'aide d'un champ extérieur.

E.Stern

Could you contrast the potential advantages of your technique with YIG film resonators?

Réponse d'Auteur

Je ne peux pas faire de comparaison n'étaient pas spécialiste de magnétisme. C'est un phénomène nouveau, mais pour le moment les inconvénients sont supérieurs aux avantages. Ceci peut changer prochainement.

Comment – E.Stern

J'ai essayé il y a plusieurs années de prospecter cette voie mais je n'ai pas réussi.

PERFORMANCE LIMITATIONS OF TWO PHASE CCD'S

V.A. Browne
Allen Clark Research Centre,
The Plessey Company Limited,
Caswell, Towcester, Northants., England.

SUMMARY

The adoption of a 2-phase technology for fabricating charge coupled devices can result in a significant reduction in system complexity due to simplified clock driving requirements. There are several different techniques available for achieving 2-phase operation and in this paper some of these will be discussed and compared.

Directionality of charge transfer in a 2-phase device is obtained by incorporating asymmetry into the potential profile under each electrode. This can be achieved in a number of different ways but the most common techniques are the use of a stepped-oxide structure, implanted barriers or a combination of both. Although the stepped-oxide technique has been most widely used so far, it is likely that the implanted barrier solution will eventually be preferred since it can be used with both surface and buried channel technologies and is directly compatible with industry standard MOS processes.

The simplification in the clock driving electronics required for 2-phase compared with 3-phase devices results partly from the reduced number of clock phases, but more significantly from their relative insensitivity to the driving waveforms. Overlapping or non-overlapping 2-phase clocks with fast or slow edges can be used and a further simplification can be achieved by holding one clock phase at an intermediate D.C. potential. This mode of operation called "1.1/2-phase clocking" results in only one clock phase being required for charge transfer. Techniques have also been developed to enable the charge injection and sensing functions to be performed using the same clock phase, thereby resulting in devices requiring only a single clock for all operations. This simple clock driving requirement results in much higher operating speeds being possible in practical signal processing systems.

In this paper the performance of 2-phase CCD's fabricated with a coplanar, N-channel, double-level polysilicon gate process will be described and compared with theoretical predictions. Factors such as charge transfer inefficiency and thermal leakage will be discussed and compared between devices fabricated with stepped-oxide, implanted barriers and offset-gate structures. A comparison between surface and buried channel devices will also be made.

1. INTRODUCTION

One of the principal problems with CCD devices has been the complex and critical drive requirements. In many cases, three or more clock phases are required to perform inputting, transfer and outputting of signal and the relative timing of these phases is usually critical. Furthermore, the large interphase capacitance of practical CCD devices can give rise to clock 'glitches' which in turn can degrade the CCD performance.

If CCD devices are to be adopted by system designers, the above shortcomings must be minimised, and preferably eliminated, even if some of the potential performance of the CCD has to be sacrificed. The goal we seek is a charge transfer system which may be reproducibly and simply set up to give an acceptable level of performance, preferably using a single phase clock.

The basic 3-phase charge transfer system has been studied by the author (BROWNE, V.A. and PERKINS, K.D., 1976) and, whilst such structures are simple in concept, they are difficult to operate. In particular they are vulnerable to clock 'glitches' which may be only a few nanoseconds in duration and which may cause the potential barriers within the device to collapse momentarily. An improvement in this situation may be realised by the fabrication of 4-phase structures whereby the charge transfer directionality is more reliably defined. However, the ideal solution is the adoption of a two phase system in which the transfer barriers are physically built into the device structure. Hence the basis of our approach to simplified CCD operation is the use of a 2-phase structure. This paper discusses the performance limitations of a number of 2-phase CCD structures when operated in both the 2-phase and the uniphase mode. Performance is compared to the equivalent 3-phase device.

2. DEVICE STRUCTURE

2-phase CCD devices have been fabricated using a coplanar, n-channel, double level polysilicon gate process. The necessary barriers to inbuild charge transfer directionality have been formed by means of either a thick oxide or an ion implant beneath the second level polysilicon, and self aligned with the first level electrodes. A typical cross section of the CCD device is shown in Fig.1. The composite 2-phase electrodes have been formed by connecting one electrode from each of the two polysilicon levels to a common busbar. Figs.2 (a) and (b) show the measured relation between surface potential and gate voltage for the two types of barrier. Note that for the thick oxide device the barrier height (represented by the separation between the two curves) increases with clock level. Conversely the shallow ion implant of the device of Fig.2 (b) behaves as an additional Q_{ss} term giving rise to a near constant shift in surface potential and a near constant barrier amplitude.

When an ion implant is used to form the barriers it is also possible to fabricate bulk channel devices using the same process schedule, with the addition of a phosphorus channel implant. Fig.2 (c) shows the typical channel potential against gate voltage curves of a bulk mode CCD. The implant barrier structure is to be preferred owing to its compatibility with both surface and bulk mode devices and owing to the good approximation to a constant barrier height which is achieved. Such a process requires the addition of just one ion implant to the existing 16K MOSRAM processes.

The barrier size has been chosen subject to the constraints of a 15 volt supply rail and a pseudo uniphase clocking system. After considering all the system tolerances, barrier sizes of 5 volts for the surface mode devices and 4 volts for the buried mode devices were chosen, corresponding to charge storage densities of $1.1 \times 10^{12}/cm^2$ and $4.0 \times 10^{11}/cm^2$ respectively. For the surface mode device in particular this represents a significant reduction in storage density compared to the equivalent 3-phase device. For many real applications whereby one clock phase is held at an intermediate D.C. potential, the storage densities of the 2-phase and 3-phase devices are similar.

An alternative structure having offset electrodes (BOWER, R.W., ZIMMERMAN, T.A. and MOHSEN, A.M., 1973) instead of the composite electrodes of Fig.1 has been considered. The system is shown schematically in Fig.3 and is seen to consist of electrodes fabricated alternately with thick oxide and ion implant barriers. Otherwise the device operation is similar to that described for Fig.1. The offset electrode system does overcome the main shortcoming of the composite electrode device by significantly reducing the element pitch. However, this is at the expense of a more complex and difficult to control process. In partic-ular the electrode dimensions of the offset device can vary with mask alignment.

In practice the real choice between offset and composite gate structures must depend upon the application to be considered. For many signal processing functions the elemental pitch of the composite gate device is not a significant problem and its well defined electrode dimensions are attractive. By contrast, for large area imaging chips the offset electrode structure has the benefit of being immune to intra-level shorts (see Fig.3).

3.   DEVICE OPERATION

3.1.   Transfer Inefficiency

All of the device structures described in Section 2 have been fabricated and the use of 2-phase and uni-phase clocking schemes have been investigated. The general reproducibility of device performance has been a notable feature. Of particular interest is the reduced sensitivity to clock waveform variations as compared to 3-phase CCD structures. The device operating characteristics are found to be largely independent of clock amplitude over the range 8-20 volts when a 2-phase clocking scheme is employed. Even devices having severe interphase electrode shorts, which give rise to grossly distorted clock wave-forms, are found to function.

A 64 stage, composite gate, implanted barrier device is illustrated in Fig.4. The device has two parallel transfer channels clocked by a common electrode set. The respective channel widths are 5 microns and 50 microns. Fig.5 (a) illustrates the operation of the two CCD delay lines, using a 12 volt overlapping 2-phase clock. The imperfect step function response is a consequence of incomplete charge transfer (trans-fer inefficiency). Clearly the transfer inefficiency of the 5 micron device is greater than that of the 50 micron device. This results from fast interface state trapping along the parallel edges of the CCD channel, the effect being proportionally larger for the 5 micron device.

Edge effects due to finite potential gradients at the electrode edges have been recognised as a limiting factor for narrow surface channel CCD devices (TOMPSETT, M.F., 1973) irrespective of the number of clock phases. The background or 'fat zero' charge is unable to communicate with fast states at the electrode edges and hence trapping of signal charge occurs. The gradient achieved at the edges of the potential wells depends upon the device geometry. The local oxidation technique offers an attractive approach to controlling the geometry of the parallel edges of the potential well by means of a self aligned field implant and thick oxide. The gradient at the transverse edges of the well is set by the electrode separ-ation, the gate dielectric thickness and the substrate doping concentration, all of which are well con-trolled also. Hence, for a given gate bias, the potential gradient at the electrode edges is well defined, and so, therefore, is the edge trapping effect.

One further factor might be expected to influence the extent of edge trapping in 2-phase CCD devices. If the receiving electrode is turned on before the sending electrode is turned off, then the charge is always pushed over a constant barrier potential so that the same channel area is covered by both the back-ground charge and the signal charge. Hence it might be expected that this "overlapping" or "push" clock mode of operation would reduce the edge trapping effects in narrow channel surface mode CCD's. Con-versely, if the sending electrode turns off before the receiving electrode turns on, then the area covered by the background charge will be smaller than that covered by the signal charge so that larger edge trapping effects may be expected for this "underlapping" or "drop" clock mode of operation.

The transfer efficiencies of the 5 micron and the 50 micron channel devices of Fig.4 have been investigated as a function of clocking mode and frequency. The net charge loss from the leading edge of the step function (Fig.5 (a)) was found to be largely independent of clocking mode for either overlapping or under-lapping 12 volt 2-phase waveforms or an 18 volt uniphase waveform. In all cases the step function was proceeded by 1024 background charge packets. Similar results were obtained for the offset electrode structure and for the composite electrode structure using thick oxide barriers. The net charge loss from the leading edge of the bulk mode device was significantly smaller owing to the reduced loss to trapping sites.

Since the surface potential curves of Fig.2 are not perfectly parallel, the transfer barriers of the surface mode device are higher when operated with overlapping drive waveforms. The resultant increase in charge density is reflected in a lower fractional charge loss for the leading pulse. Fig.5 (b) shows the measured fractional charge loss per transfer (transfer inefficiency) for the various operating modes as a function of frequency.

Two important points were noted in taking these measurements:

        (i)   In the uniphase mode all devices were designed to operate in the complete transfer mode using a 15 volt transfer clock. However, for the surface mode devices it was found necessary to use a minimum of 17 volts amplitude before the complete transfer mode was achieved. With the 18 volt transfer

clock employed for the above measurements the adjustment of the D.C. clock phase was found to be critical if the complete transfer mode of operation was to be achieved. In contrast, the bulk mode device operated satisfactorily using a 15 volt uniphase clock and also had a wide tolerance (±2 volts) on the D.C. phase setting, within which range the performance was found to be constant.

Clearly the surface mode device did not behave as predicted by its measured channel potential curves when using the uniphase clock. On the other hand, the measured channel potentials enabled all other aspects of the device performance to be accurately predicted. It is likely that the discrepancies in the uniphase mode arise from the existence of spurious potential barriers between the CCD electrodes. Fig.6 shows how such barriers can exist in the overlapping electrode 2-phase CCD. Neither the ion implant barrier nor the oxide beneath the barrier gate are truly self aligned to the first polysilicon electrode edges. In reality both the implant and oxidation reach beneath the electrode edges by the relevant diffusion length giving rise to spurious, but stable, barriers. In the case of 2-phase operation, or 2-phase bulk mode operation, the larger fringing fields eliminate the spurious barriers so that complete transfer mode operation is achieved.

(ii) The predicted reduction of edge effect losses when using two phase overlapping waveforms was not observed, even in the case of the narrow channel devices where performance is dominated by edge effects. It may be that the small remaining fraction of charge which is slow to transfer from the sending electrode suffers edge effects in the normal manner as the clock turns off. An accurate modelling of charge transfer in 2-phase CCD's is needed if these effects are to be fully understood.

At high clock rates the fractional loss per transfer is seen to increase owing to the finite mobility of the signal charges. The rate at which free transfer losses increase at high frequencies has been found to be generally less than predicted by the established theory (CARNES, J.E., KOSONOCKY, W.F. and RAMBERG, E.G., 1972).

3.2.   Thermal Leakage

Thermal generation of charge has been measured at typically 10 nA/cm^2 at 25°C. However, localised defects can limit the saturation time of some storage cells to typically 1-5 seconds in surface mode devices and 0.1 to 0.5 seconds in bulk mode devices. The origin of thermal charge from three sources has to be considered in the coplanar CCD viz:

(i)   Direct and indirect band to band generation in the non-defect cells. In the case of the barrier electrodes and the 'off' storage electrode, where the depletion layer width is small, this will generally be dominated by carriers hopping via mid-band interface states. In the case of the 'on' storage electrode, where the depletion width can be large, bulk generation dominates.

(ii)   Indirect generation in cells having crystallographic defects which can act as nucleation sites for impurity ions. In coplanar structures careful processing is required to avoid stacking faults in the field regions which can extend into the active channel. Even in wafers processed with a zero density of stacking faults, leakage 'spikes' are found to occur as the second order defects nucleate impurity ions.

(iii) Leakage current arriving from the field regions. Although the surface band-bending is small in the field regions, a finite depletion layer exists, especially beneath positively biased rails. The use of collection channels of the type used for antiblooming protection in vidicons can help eliminate the possibility of spurious injection from outside the active channel. Alternatively an n+ collection ring may be employed.

In 2-phase CCD's, which have an approximately constant barrier amplitude, the storage time capability may be enhanced by storing signal with both clocks in the off condition. This is especially attractive in the surface mode device whereby the depletion volume beneath the 'on' storage gate is reduced. Also inversion charge is then collected beneath both storage gates resulting in a reduced surface generation current since the occupied surface states can no longer act as generation levels.

3.3.   Charge Storage Density

Surface Mode

The composite gate device of Fig.4 has been fabricated with 9 volt barriers, so that 12 volt 2-phase operation gives rise to a charge storage density of 2×10^{12} charges/sq.cm. This compares well with the equivalent 3-phase device fabricated on the same process and using a 3-phase, twelve volt clock.

When uniphase mode operation is required, however, the barrier height must be limited to approximately 5 volts if complete transfer operation is to be achieved. Thus the performance of the 2-phase device has to be compromised when uniphase operation is required. This is analogous to operating the 3-phase CCD in the 2½-phase mode.

Bulk Mode

The charge storage density of the bulk mode device will generally be limited by the capacity of the buried channel. Consequently the same storage density is achieved irrespective of whether 2-phase or uniphase operation is considered. The bulk mode devices described in this paper had a barrier height of 4 volts and a storage density of 4×10^{11} charges/sq.cm.

4.   CONCLUSIONS

2-phase surface and bulk mode CCD's can be fabricated using the two level polysilicon RAM processes, with the addition of just one implant for surface mode operation and two ion implants for bulk mode operation.

Operation of 2-phase CCD's is more predictable than other forms of CCD owing to the inbuilt barriers which ensure the existence of discrete potential wells at all times. Moreover, operation of the bulk mode device may be simplified by the use of a single phase transfer clock. This simplification involves no compromise in performance. For the surface mode device, however, the barrier height has to be limited if single phase operation is intended, with the result that performance similar to the 3-phase CCD operated in the 2½-phase mode is achieved. One of the more serious problems to be addressed when using uniphase surface mode operation is the existence of spurious potential barriers which can degrade performance. The reduced fringing fields which exist between electrodes during surface uniphase operation can be too small to eliminate such barriers.

5.   REFERENCES

BROWNE, V.A., PERKINS, K.D., Feb. 1976, "A Non-overlapping Gate Charge Coupling Technology for Serial Memory and Signal Processing Applications", I.E.E. Journal of Solid State Circuits, Vol.SC-11.

BOWER, R.W., ZIMMERMAN, T.A., MOHSEN, A.M., 1973, "A High Density Overlapping Gate Charge Coupled Device Array", Proceedings of Intl. Electron Devices Meeting, Washington.

TOMPSETT, M.F., Jan. 1973, "The Quantitative Effects of Interface States on the Performance of Charge Coupled Devices", I.E.E.E. Trans. on Electron Devices, ED-20, p45.

CARNES, J.E., KOSONOCKY, W.F., RAMBERG, E.G., June 1972, "Free Charge Transfer in Charge Coupled Devices", I.E.E. Trans on Electron Devices.

McCAUGHAN, D., HARP, G., 1976, "Phase Referred Input - A Simple New Linear CCD Input Method", Electronics Letters 12, 682.

GOODING, J.N.,   This Conference.
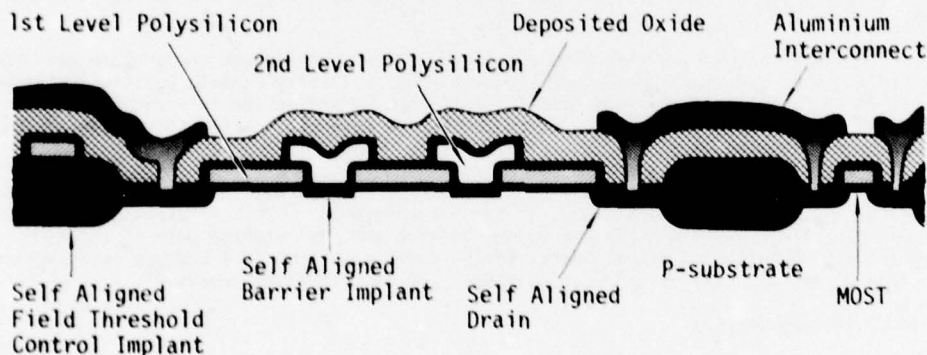
FIG. 1

SCHEMATIC CROSS SECTION OF THE COPLANAR, COMPOSITE GATE CCD STRUCTURE, SHOWING HOW THE FIELD OXIDE, FIELD IMPLANT, BARRIER IMPLANT AND DIFFUSIONS ARE SELF REGISTERED.

Polysilicon 1 + 1000 Å Oxide

Polysilicon 2 + 3000 Å Oxide

Surface Potential (Volts)

Gate Volts

(a)

Polysilicon 1 + 1000 Å Oxide

Polysilicon 2 + 1500 Å Oxide
+ 10^{12} Boron/cm²

Surface Potential (Volts)

Gate Volts

(b)

Polysilicon 1

Polysilicon 2 + 8 x 10^{11}/cm² Boron

Channel Potential (Volts)
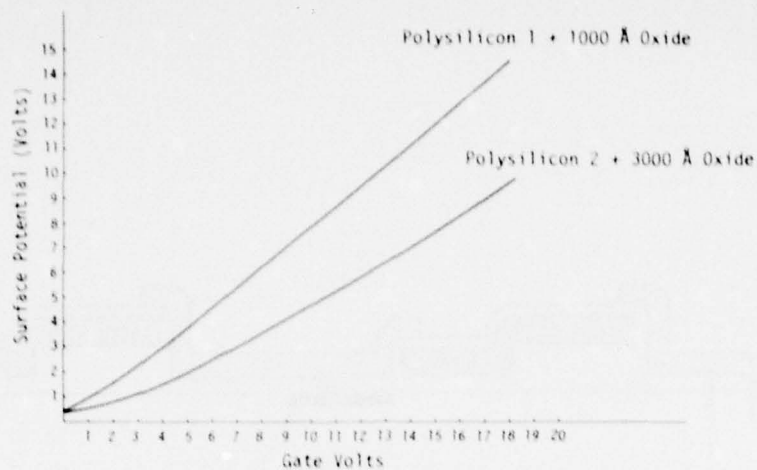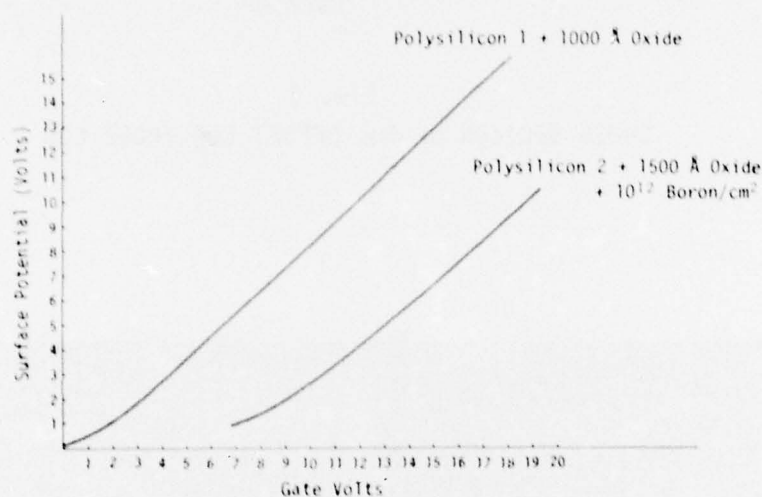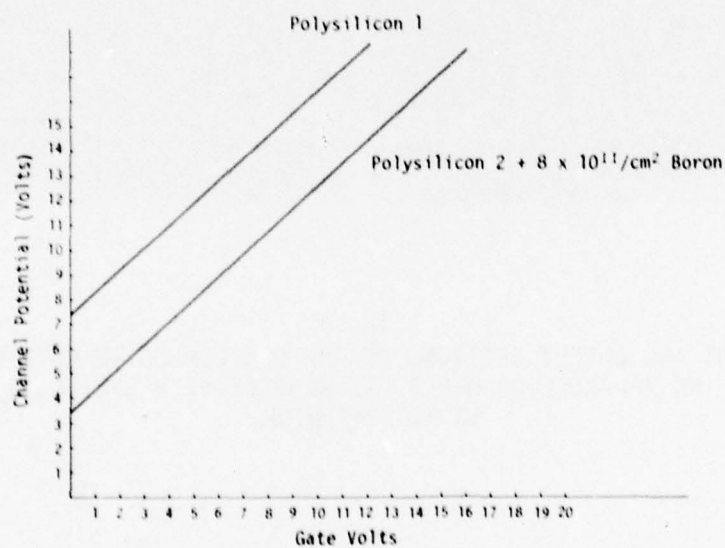
Gate Volts

(c)

FIG. 2
POTENTIAL OF THE CCD CHANNEL AS A FUNCTION OF APPLIED
GATE VOLTAGE FOR (a) THICK OXIDE BARRIERS (b) ION
IMPLANT BARRIERS AND (c) BULK MODE OPERATION WITH
ION IMPLANT BARRIERS.

2nd Level
Polysilicon

Self Aligned
Directionality
Barrier

1st Level
Polysilicon
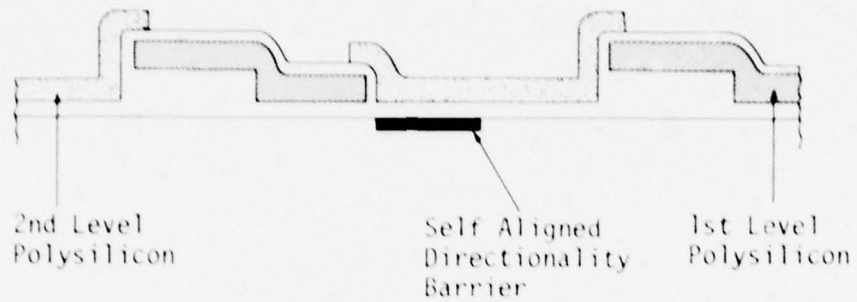
FIG. 3

CROSS SECTION OF THE OFFSET ELECTRODE CCD
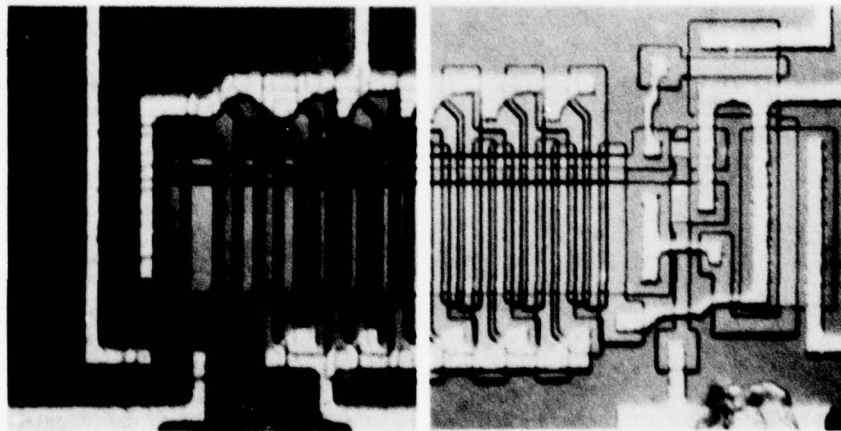
FIG. 4

INPUT AND OUTPUT SECTIONS OF THE 2 Ø COMPOSITE GATE CCD.
THE TWO PARALLEL CHANNELS ARE RESPECTIVELY 5 MICRONS AND
50 MICRONS WIDE.
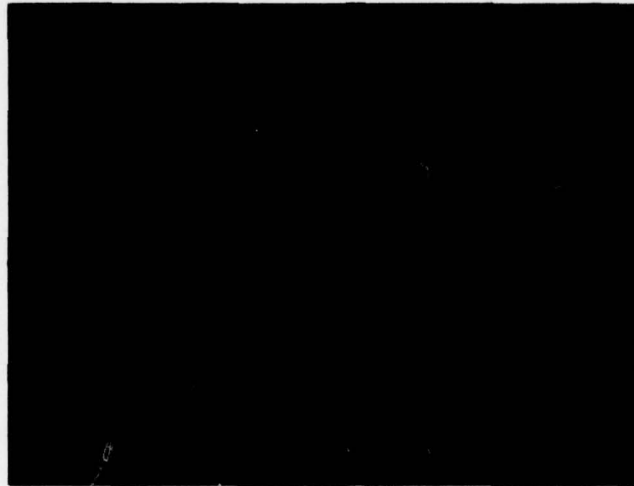
(a) STEP FUNCTION RESPONSE OF THE TWO CCDs OF FIG. 4.
    UPPER TRACE = 50 MICRON WIDE CHANNEL
    LOWER TRACE = 5 MICRON WIDE CHANNEL
    VERT = 2v/div and 0.5 v/div. HORIZ = 100 5/div.

o Surface Mode. 5 Micron Channel 2∅ Underlapping and Uni ∅
□ Surface Mode. 5 Micron Channel 2∅ Overlapping
△ Surface Mode. 50 Micron Channel 2∅ Underlapping and Uni ∅
× Surface Mode. 50 Micron Channel 2∅ Overlapping
● Bulk Mode. Uni ∅
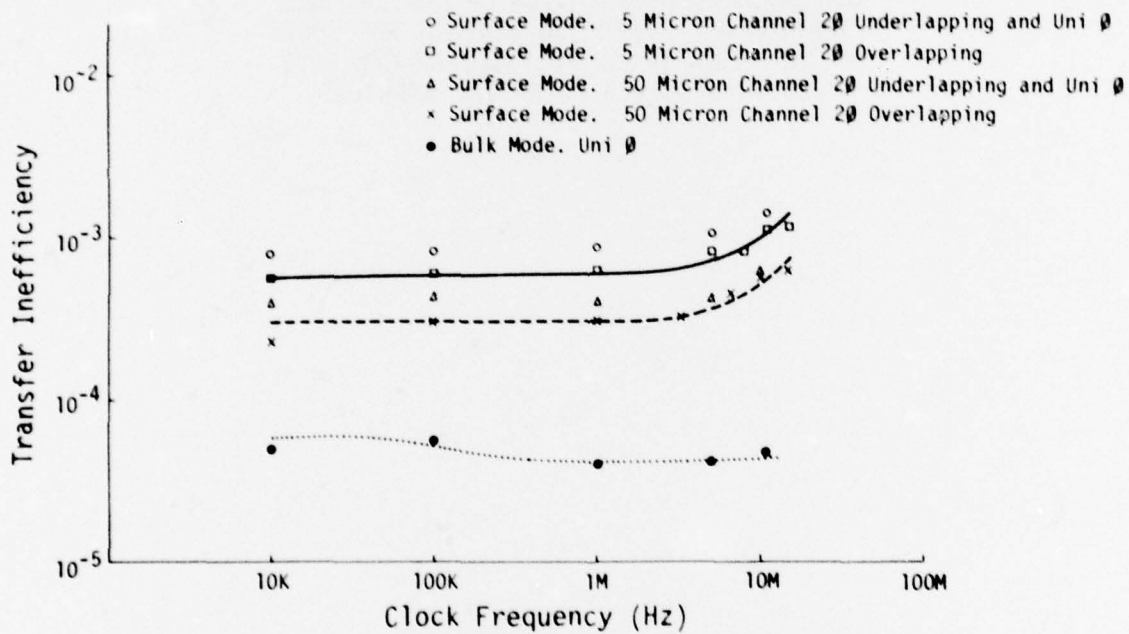
(b) TRANSFER EFFICIENCY OF THE IMPLANTED BARRIER
    COMPOSITE GATE CCD AS A FUNCTION OF OPERATING
    MODE AND FREQUENCY.

V_{CLOCK} = 12v (2∅) OR 18v (UNI - ∅)

V_{BB} = -5v

FIG. 5
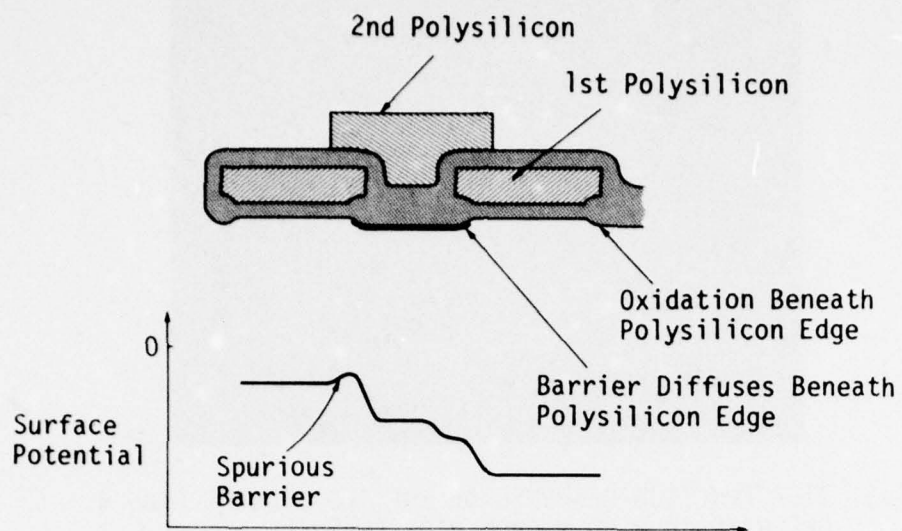
FIG. 6
SCHEMATIC REPRESENTATION OF HOW LATERAL
OXIDATION AND DIFFUSION BENEATH THE EDGES
OF THE FIRST LEVEL POLYSILICON CAN GIVE
RISE TO SPURIOUS POTENTIAL BARRIERS.

DISCUSSION

Roberts

How do you define the thermal leakage times you specify: charge packet half-life, 10%, or 90% or what?

Author's Reply

Times stated are for saturation.

In practical case, a portion of signal range (say 20%) must be allocated to leakage charge. In this case the useful storage time would be 1/5 of the stated times.

J.J.Stapleton

Which of your 2-phase CCD techniques was better for driving the voltage-dependent capacitances of the clock lines?

Author's Reply

The capacitance which the clock driver sees is dominated by poly 2/poly 1 overlap capacitance. This is not voltage-dependent. This voltage may be about 100 pF and thus prevent driving problems. This will be discussed further in paper 3.9.

The voltage-dependent capacitance referred to in paper 3.9-5 is more relevant to nodes carrying analogue signals than to clock nodes.

THE DESIGN AND DEVELOPMENT OF CCD PROGRAMMABLE TRANSVERSAL FILTERS AND CORRELATORS

J. Mavor   and   P.B. Denyer

Electrical Engineering Department,   Wolfson Microelectronics Liaison Unit,
University of Edinburgh
Mayfield Road, Edinburgh, EH9 3JL
Scotland

SUMMARY

This paper examines the potential of integrated circuit-based convolvers for baseband operation. The emphasis is on compact processor development with an eventual aim of a single chip, monolithic design suitable for a variety of signal processing applications. Before hardware developments are treated, trade-offs of serial and parallel configurations of convolver are considered together with their respective component requirements. It appears that for complete integration of a convolver module, the parallel approach is preferred. Following a description of a Deltic serial processor (having a time-bandwidth product of 500) a study is made of the circuit elements of a particular realisation of an integrated convolver in CCD/MOS technology. All aspects of the design are considered and the description is focussed around a prototype 64-stage convolver. The future of this approach is discussed and thought to be capable of 256 convolution points on a single chip. Several such units could be cascaded giving this integrated approach an enormous potential in applications where space and power are crucial, and high computing precision is not essential.

1.   INTRODUCTION

The development of compact, lightweight, low-power hardware to perform the real-time correlation/convolution of two discrete signals permits the realisation of sophisticated signal processing functions, such as matched filtering and spectrum analysis for a variety of applications. Recent advances in microelectronic components are central to these developments and include progress in analogue as well as digital devices. Now, in addition to conventional digital elements for performing discrete-time sampled, signal processing, charge-coupled devices (CCD's) (Boyle, W.S. and Smith, G.E., 1970) are available. These provide primarily high density analogue/digital storage in various configurations, and are compatible with MOS circuitry for performing ancilliary functions. The CCD/MOS approach looks extremely promising for realising convolvers with performance attributes suitable for many areas including; sonar, communications and instrumentation, particularly for mobile equipments.

The advent of the microprocessor and LSI memory means that extremely powerful and versatile filtering systems can be produced by controlling the convolver reference channel. Such an arrangement is extremely important in the signal processing context as any desired filtering function can be realised simply by programming its impulse response. Interfaced with a microprocessor (Copeland, M.A., 1977) this system could provide 'intelligent' adaption of the filter response to optimise performance, or for restricted applications, a single filter may be multiplexed to implement many different signal processing functions.

The crossconvolver can be realised using various architectures based upon combinations of serial and parallel forms. Each configuration is likely to be suited to a particular hardware arrangement and dependent upon the overall system parameters. In this paper, the emphasis is on solid-state convolvers based on silicon integrated circuit technology. The trade-offs between, and limitations of, serial and parallel forms of processor are discussed with particular reference to integrated circuit implementation with a minimum package count. Practical aspects of both forms of convolver realisation are discussed with reference to prototype hardware realisations.

2.   PRINCIPLES

2.1 Convolution

A filter having an impulse response sequence, r, of N elements will respond to an L element signal sequence, s, with an (L + N) element output sequence, given by the convolution sum (Squire, W.D. et. al., 1969):

$$(s * r) (m) = \sum_{n=0}^{N-1} r_n \cdot s_{m-n} \qquad (2.1.1)$$

The calculation requires N multiplications for each convolved point (one multiplication for each reference element) and, thus, a total of (L + N)N multiplications for the complete convolution sequence.

2.2 Serial vs. Parallel

Large computers can be used to programme and carry out these computations very easily, but the demand for dedicated, compact, hardware realisations is an important and developing applications area.

All convolution systems must provide some memory register for storing the reference sequence and the present and previous, N, signal samples, $s_{m-n}\ldots\ldots\ldots s_m$; a facility for multiplying together individual samples of the two sequences; and a method of shifting the sequences with respect to each other, by one element, after each calculation of the convolution sum.

One common approach for realising the convolution function is shown in principle in Fig.1(a), and employs a single multiplier to execute the N multiplications at each convolved point in series. Thus for real-time convolution, the multiplier must operate at N times the signal sample frequency; similarly, the data contained in the signal and reference registers must be recirculated at that rate. After each

recirculation, the oldest signal sample is replaced by a new input sample and the next convolution point is obtained by a further recirculation.

As an alternative to the serial system, real-time computation of the convolution sum may be achieved at a processing rate equal to the signal sampling rate by the effective *parallel* processing of the required multiplications. This may be realised in an N-stage, programmable transversal filter of the type shown in Fig.1(b). Here the hardware requirement is increased by N-1 multipliers, but for a given signal bandwidth the speed requirement is reduced by a factor of N. Note also that the equivalent number of signal sample shifts, or transfers, is also reduced by a factor, N.

Many other correlator architectures are possible usually involving one or more serial/parallel configurations to optimise the parameters of the system. In many cases the speed of the multiplication hardware limits the operating bandwidth and the system has to be configured with multiplexed or parallel channels to achieve performance at the expense of duplicated multiplication.

2.3 Analogue vs. Digital Processing

Digital convolvers based on conventional digital components are widely available but suffer from the disadvantage that when they are used to process analogue signals - as is very often the case in signal processing - A/D conversion is required. Their speed is limited essentially by the rate of the digital multipliers used, and these are likely to be largely responsible for the high power taken by such a processor. Furthermore, digital multiplication is costly in terms of hardware, requiring currently multi-chip systems for anything above four-bit accuracy. For higher packing density, analogue multipliers and summers can be used but these, of course, suffer from the inacuracies of analogue multiplication and thus impose an upper limit on useful resolution. However, the distinct advantage of a digital convolver - in common with all digital systems - is that the accuracy of computation, which is a direct function of the digital word size, is essentially infinitely expandable (but naturally limited by the components used). This precision is obtained at the expense of power and system size, but for some applications is of major importance.

Analogue processors of all kinds suffer from the disadvantages of drift, sensitivity and difficult circuit design. Convolver realisations based on analogue techniques are also subject to the restrictions of the usually modest dynamic range of analogue-analogue signal multipliers. However, where these problems can be contained they have the attributes of a *more direct* realisation of the convolution function, in comparison with digital approaches; thus advantages may arise which outweigh the inherent problems for a particular application.

In realising an analogue processor, analogue memory could be provided by MOS capacitor storage, however, by using charge-coupled device memory then the shift facility is provided inherently. Thus, CCD-based convolvers offer potentially several salient advantages over all-digital techniques. Their application is substantially limited to parallel, or transversal filter, realisations of convolver because of the difficulty of recirculating sampled analogue data in an analogue-mode CCD. Charge-transfer efficiency problems degrade the data after many recirculations and make serial convolver forms essentially impossible to implement, even with compensation for charge dispersion.

The provision of an analogue reference memory is an additional problem to be faced; conventional CCD/MOS processes do not allow long term storage of analogue information and some means of refreshing must be provided. This may not be such a severe disadvantage at the systems level, however, as the reference information may often be held in digital form in some memory associated with a central control unit (possibly a microprocessor). Refreshing the reference can then take on a 'direct memory access' (DMA) priority without interfering with other control functions. Thus a dedicated non-volatile memory block, always associated with a given analogue convolver is not necessarily implicated.

A further convolver formation is to have an analogue signal input and a digital reference. The multiplier requirement is relaxed where a single bit reference is suitable and, consequently, the overall hardware complexity of the processor thereby becomes reduced; analogue multipliers are, however, implied. Monolithic correlators of this type have been reported (Herrmann, E.P. et. al., 1976) in CCD/MOS technology. In order to simulate a multi-bit reference, several units can be run in parallel or one longer unit time-shared. As this approach has been discussed elsewhere, this paper will not consider it further but concentrate on analogue-analogue signal convolvers where the highest computing density and most interesting possibilities arise.

2.4 Comparison of Realisations

It is clear from the preceding discussion that if high precision is required then a digital processor is essential. For low bandwidth applications (< 20 kHz for a time-bandwidth of 1000) then a serial convolver offers many advantages. However, where dynamic ranges of 40 - 60 dB are suitable, or a high speed processor is required, then a parallel approach based upon CCD/MOS devices may be a possibility. Therefore, as digital component development progresses, the boundary between the various approaches will reduce the efficacy of analogue approaches, although analogue-analogue convolvers will always provide the maximum packing density in parallel implementations.

Table 1 summarises the features of the various approaches: performance figures are not included because of the difficulty of specifying hybrid systems and comparing performance parameters. Many trade-offs are evident from the Table but three particular cases appear most interesting:

(a) The parallel, digital convolver represents the best combination of speed and accuracy at the expense of power and size;

(b) The parallel, analogue convolver represents the best combination of speed and packing density, at the expense of accuracy, and;

(c) The serial, digital convolver represents the best combination of packing density and accuracy for low frequency applications.

Of these three approaches, the serial digital convolver and the parallel analogue convolver are discussed further in this paper.

Table I: PERFORMANCE COMPARISON OF DIFFERENT CONVOLVER REALISATIONS

Convolver Type		Speed	Packing Density	Accuracy	Features
Configuration	Mode				
SERIAL	DIGITAL	Medium-Slow (Trade-off with packing density)	Medium-High (Trade-off with accuracy)	To any requirement (with digital multiplication)	Permanent, signal and reference memory
	ANALOGUE and DIGITAL			Limited by analogue multipliers	Restricted by recirculation of analogue data
	ANALOGUE				
PARALLEL	DIGITAL	Fast	Low	To any requirement (with digital multiplication)	Permanent, signal and reference memory
	ANALOGUE and DIGITAL		Medium	Limited by analogue multipliers	Permanent reference memory
	ANALOGUE		High		Restricted by analogue reference decay

3. SERIAL CONVOLVER IMPLEMENTATION

In section 2.2 the serial convolver principle was described with reference to Fig.1(a). In this case the signal is time compressed, or expanded in bandwidth, by a factor N to permit a single, fast multiplier element to perform the required (L + N)N multiplications in a time equal to N low data rate multiplications.

A processor based on the DELTIC* configuration using CCD digital storage and digital multipliers has been reported (Mavor, J. et. al., 1977), and is illustrated in Fig.2. The storage was organised in a format of 1024 bytes x 8 bits, and an 8 x 8 bit multiplier was used in the prototype. Although the maximum memory clock rate was 3 MHz, during test the convolver was operated with a 500 Hz input signal bandwidth only requiring a clock rate of 1.025 MHz. The processor can therefore process 1000 samples of a waveform of 1 s duration and bandwidth 500 Hz; this corresponds to a time-bandwidth product of 500. Whilst the *principle* of the technique has been established, clearly, using state-of-the-art multipliers full sonar bandwidth operation must ultimately be achievable at time-bandwidths extendable in blocks of 1000.

4. PARALLEL CONVOLVER IMPLEMENTATION

The remainder of the paper is devoted to a systematic study of the design philosophy regarding a particular realisation of an integrated, parallel convolver based on the principles of section 2.2. Because a single chip design only appears achievable using analogue-mode CCD/MOS, and because the convolver in this technology becomes an *active* component, then the design problem is very involved (Harp, J.G. et. al., 1975) (Bosshart, P., 1976). In this section, details of many of these considerations are outlined, pertaining to the convolver signal input, CCD tapped delay line, the four-quadrant multipliers and a programmable reference memory. Although many other approaches to a single chip convolver are possible, it is thought that the majority of the discussion will be common to other realisations based on CCD/MOS technology.

4.1 MOS Design Techniques

In this section, the requirements for the MOS circuitry necessary to implement a convolver, when combined with a CCD delay line, are outlined. Later work will present the limiting processes and relevant equations that apply in the design of the integrated circuit.

MOS linear circuitry concepts have not been fully exploited to date for any purpose because of the problems associated with this approach. Hitherto, problems of noise, drift and low transconductance

*Delay Line Time Compressor

values (gain) have made their application to linear processing very restricted, in comparison to bipolar transistors. Although bipolar types have many useful operating parameter values, they do suffer from being relatively large in area. In the convolver design, where the object is to pack as many processing cells as possible on the chip, MOS transistors would have the advantage of small size. Thus in the study leading up to the prototype convolver design, detailed later in the paper, the adverse features of MOS structures were studied in their required configurations to see if any improvements could be innovated at the circuit level.

Fortunately, in the CCD context, MOS linear circuits can be made to appear to be high performance by adopting an elegant technique. The sampled-data nature of CCD operation imposes on the MOS circuitry its time quantised data restriction. Thus the MOS circuitry serving, say, the outputs of a convolver CCD delay line, is required to be active for the time when the data is valid and then is not used. During these regular periods in which the data is being transferred in the CCD (say about half of a clock period for a $2 - \phi$ device) any drift in the non-sampled MOS circuitry can be assessed and then eliminated by an appropriate correction. This could involve, for example, storing an error signal caused by transistor drift on a capacitor (store) so that it may be subtracted from a subsequent, new signal sample during the next valid data time. As the error correction process could be maintained at the CCD clock p.r.f. the 'chopping' frequency could be very high and should be extremely effective in minimising any drift that would otherwise occur. It is also very useful as an approach because it could be used not only for correcting say the drift in a monolithic MOS transistor summing amplifier, but also for reducing d.c. drift in the bias levels of the CCD tap amplifiers and multipliers.

A further technique has been proposed (MacLennan, D.J. and Mavor, J., 1975) for improving the input/output linearity of CCD/MOS circuits. Without this technique the gain and distortion in such circuits would otherwise be ill-defined, and lead to specification problems of the convolver system. The technique uses a single, high gain, operational amplifier for improving the linearity by including the otherwise non-linear input circuit in the forward-gain loop of a voltage follower. Further details of this approach to linearising the input/output transfer function are given in the following sections.

The basic operations of a parallel convolver for monolithic implementation have been discussed previously and are summarised in Fig.1(b). In order to achieve the necessary density of functions in an integrated circuit, all the operations such as, delay, multiply and sum, must be obtained within a narrow strip on the surface of the silicon substrate. In practice, a major problem in layout is the four-quadrant multipliers. These should be fabricated at a pitch corresponding to the CCD delay line cell length, or in a multiple of it, for a neat device topology. In section 4.4, considerations of an MOS multiplier structure are detailed together with an error analysis for their performance.

In the remainder of this section other details of the MOS circuit design requirements are given, including a technique for storing and presenting the reference signal to the multipliers, and details of the summing circuitry. This work forms a unified basis for a fully integrated convolver design in CCD/MOS technology.

4.2 Signal Register

A prime requirement of the transversal filter realisation of Fig.1(b) is a periodically tapped, analogue delay line. Such sampled-signal delay functions are readily implemented using compact, low power CCD registers. These provide naturally the sequential delay of analogue signal samples, in the form of isolated charge packets, via a series of closely spaced controlling electrodes. The charge packets, which are amplitude modulated to contain signal information, are transferred along the register by driving the electrodes with m-phase clock waveforms (typically $m = 2, 3, 4$). Clearly, the time delay between samples is directly controlled by the clock frequency. An additional requirement of the tapped delay line, however, is that the information contained in each sample be periodically sensed whilst leaving the charge packet isolated and free to transfer along the register.

A non-destructive sensing scheme (Denyer, P.B. and Mavor, J., 1977) which may be implemented in any CCD technology is illustrated in Fig.3. A conventional clocked electrode in the CCD structure is replaced by a 'floating gate' (reset) sense (FGR) electrode. This electrode is firstly reset to a bias voltage whilst there is no signal charge underneath it. It is then isolated and any charge subsequently injected (as a result of the CCD clocking action) causes a related change in electrode potential. The signal may be sensed via an MOS transistor amplifier to provide a voltage output at a low impedance.

Although the intrinsic transfer function may be complex the linearity and gain of this structure may be well defined by applying a feedback linearisation scheme (MacLennan, D.J. and Mavor, J., 1975) at the register input, as shown in Fig.4. In this case an FGR tap is used to meter the input charge level which is then corrected via a differential amplifier until the desired tap output signal is obtained. The charge packet is isolated and transferred along the CCD delay line resulting in identical output signals at all subsequent taps, which are replicas of the input metering tap.

A typical delayed tap output resulting from a chirp input signal is shown in Fig.5. Note that the output signal is 'chopped' and the true signal information appears for only a fraction of the complete clock period due to the reset-isolate-inject charge sensing sequence.

Considering errors within the structure, the CCD tapped delay element has three fundamental disadvantages:

(a) Transfer inefficiency - A small fraction of each charge packet is left behind at each transfer, causing cumulative signal degradation along the register. For an inefficiency, ϵ, per tapped stage, then the output at a non-destructive tap, n, at time, m, is:

$$\sum_{r=0}^{r=(m-n)} \frac{(n + r)!}{n!\,r!} \epsilon^r (1 - \epsilon)^n s_{m-n-r}$$

In the frequency domain this has the effect of imposing a low-pass filter characteristic the severity of which increases with increasing cumulative transfer inefficiency.

(b) Dark Current - As signal charge packets travel down the CCD they are subject to an additive error due to minority carriers (Tasch, A.F. et. al., 1973), which are thermally generated within the semiconductor. For a uniformly clocked CCD, the net effect is a steady increase in signal charge with time. As the effect is essentially linear there results a constant offset error at tap, n, of:

$$\hat{V}_{sig} \frac{J_D A_C T_t n}{\hat{Q}_{sig}} \text{ volts}$$

where, $\hat{V}_{sig}$ is the change in tap output potential corresponding to the peak charge capacity, $\hat{Q}_{sig}$ J_D is the dark current density, A_C is the average CCD cell area and T_t is the time delay between taps.

(c) Amplifier mismatching - Although the CCD transfer function can be linearised at the input prior launching the signal charge packet the ultimate accuracy of the tapped delay line requires that the characteristics of subsequent tap amplifiers match those of the input metering tap. Such mismatches may be caused by random gain and threshold variations in the MOS transistors. It is possible, however, to design the buffer amplifiers such that these effects result only in quiescent offset errors. It is shown later that these errors may be cancelled at the output.

4.3 Reference Register

As the CCD signal register provides the necessary time-shift process a static reference register, permanently driving one set of multiplier ports, is sufficient. The most compact analogue reference store consists of a series of capacitors followed by buffer amplifiers, if necessary to drive the multiplier circuitry. Many reference elements may be loaded from a common reference input busbar using on-chip digital multiplex techniques, or by reading the reference serially into a CCD register and then transferring the reference set in parallel to the capacitor array. The former technique, illustrated in Fig.6, is preferred for two reasons:

(a) Real-time updating permits feedback linearisation of each reference sample to be implemented right up to the appropriate multiplier port, and

(b) by making the multiplex circuitry addressable individual reference addresses may be selected randomly as well as sequentially, which may be of great use in systems requiring permuted data (Prime Transforms for example).

The major limitation of the capacitively-held reference is information decay due to charge leakage. The dominant contribution toward leakage current occurs at the diffusion terminal of the sampling transistor, and its magnitude depends upon the characteristics of the diffused diode in a complex manner. Assuming, for simplicity, a constant leakage current, I_L, then the stored value, r, decays such that:

$$r \rightarrow r - \frac{I_L}{C} (t - t_r) \tag{4.3.1}$$

where C is the holding capacitance and t_r is the last refresh time.

Capacitive breakthrough of the sample waveform onto the hold capacitor via gate overlap capacitance is a second-order effect. The voltage level at which the sample transistor cuts off, and thus the magnitude of the breakthrough, is a function of the voltage stored. Assuming a linear buffer amplifier characteristic, the net effect is a quiescent offset combined with an apparent reduction in gain:

$$r \rightarrow (V_o - r_o - V_T) \frac{C_f}{C_f + C_h} + \frac{(1 - C_f)}{C_f + C_h} A_v r \tag{4.3.2}$$

where, r_o is the reference zero level, V_o is the off potential of the sampling waveform, V_T is the threshold voltage of the sampling transistor, A_v is the small-signal voltage gain of the buffer amplifier and C_f and C_h are the feedthrough and hold capacitances, respectively.

4.4 Multipliers

For general purpose signal processing four-quadrant multiplication of the signal and reference voltages is required at each convolution point. An economical multiplication technique is based upon the essentially linear transconductance of an MOS transistor operating in the non-saturation region. A first-order expression for the drain current of such a transistor, shown in Fig.7, is given by:

$$I_D = \beta_m \left([V_{GS} - V_T]V_{DS} - \frac{V_{DS}^2}{2} \right) \tag{4.4.1}$$

from which the change in drain current obtained upon disturbing the gate voltage by an amount V_{GS}' is

$$\Delta I_D = \beta_m V_{DS} V_{GS}' \tag{4.4.2}$$

It is worthwhile noting that V_{GS}' may take either sign and that equation (4.4.2) is equally valid if

V_{DS}, as defined in Fig.7, changes sign and the drain and source terminals become transposed[*]. Thus the desired four-quadrant multiplication may be obtained by applying the operand voltages

(a) as a constant v_r with respect to some fixed 'zero' level V_{ro} at one diffused terminal, and

(b) as a change in gate voltage v_s with respect to some 'zero' level V_{so},

and by sensing the corresponding change in drain current.

Previously the difference voltage, v_s, has been synthesised (Harp, J.G. et. al., 1975) by using two identical transistors with a common diffused terminal (v_r); the gate of the second transistor being driven with the 'zero' voltage V_{so}. The accuracy of this method is however critically dependent upon good transistor matching.

An alternative approach using a single MOS transistor is feasible where the reference voltage (v_r) is constant, as in this case. Here the undesired terms in equn.(4.4.1) also become static, representing a constant offset at the output. Changes in the gate (signal) voltage now result in product changes at the output with respect to the (reference dependent) zero level. Where necessary a defined output zero level may be obtained by alternately applying a signal zero (V_{so}) to the transistor gate to obtain the output zero current:

$$I_o = \beta_m \left([V_{so} - V_T] v_r - \frac{v_r^2}{2} \right) \qquad (4.4.3)$$

and then subtracting this value from subsequent 'signal plus zero' samples. The alternate signal zero may be simply obtained by tapping the CCD register at alternate stages and multiplexing signal and zero samples at the input.

Thus a form of chopper stabilisation may be applied at the output to set and hold a zero level during the reset period; and output, subsequently, the product signal as a change in this level. Not only does this zero multiplexing technique remove the transistor matching requirement (a time-multiplexed transistor is perfectly matched to itself) but the alternate zero levels also contain information on all quiescent offset errors in the signal port (due to tap amplifier mismatching and threshold voltage variations in the multiplier transistor), which are thus also removed at the output.

Two possible errors arise when using this transconductance multiplication technique:

(a) The changes in drain current may modulate the reference voltage, v_r, due to the finite output admittance of the reference amplifiers. If the reference value has been set (using a feedback linearisation scheme) during the signal zero phase then the product error due to a reference amplifier output admittance of g_o is given by:

$$sr \rightarrow sr - s^2 r \frac{\beta_m}{g_o} \qquad (4.4.4)$$

where the error factor β_m/g_o may be typically 1%.

(b) Individual, 'matched' transistor gain factors β_m may vary within a given integrated circuit by as much as 5% and between different circuits on the same slice by 20%. The variation between devices may be trimmed out at individual summing amplifiers but there still remains a random product gain variation, $\delta\beta/\beta$, between convolution points on any given device. The statistics of this variation are obviously process dependent but the effect may be minimised by using multiplication transistors of large area.

4.5 Summing

The provision of the multiplier reference voltage, V_{ro}, in addition to the summation of the product currents may be conveniently realised using a single operational amplifier and feedback summing resistor as shown in Fig.8. Note the single current summing bus which again allows a compact circuit topology.

The removal of the alternate zero current level may be easily accomplished at this stage by programming a current source/sink during alternate periods to remove the zero current level from the summing busbar. The removal of this quiescent current before the summing resistor enables the wanted output signal to occupy the full dynamic range of the summing amplifier rather than a restricted portion of it, thus enhancing the noise figure of the complete device and allowing the output gain to be varied without a corresponding d.c. shift.

Generally, the summing amplifier is the limiting factor in terms of system speed in that its output must settle within the time period for which the signal is true; for an m-phase device the settling time must therefore be $T_c.s/m$; where T_c is the CCD clock period. If a signal sample-hold facility is made available on-chip prior to the multiplier gate, this settling time may be extended to $T_c.s$. Although the intrinsic CCD/MOS circuitry may be designed to operate at clock frequencies up to 10 MHz, it is difficult to achieve sampling rates at the summing amplifier much above 1 MHz without resorting to bulky, high power amplifiers which make mobile, lightweight applications less attractive.

An operational amplifier design for this purpose in MOS technology has been reported, (Weste, N. and Mavor, J., 1977) using single channel transistors. It was fabricated with an N-channel CCD process and gave satisfactory performance to above 1 MHz. A chopper stabilisation technique, similar to that outlined

---

[*] Because the drain and source terminals remain at a constant potential there is no additional modulation of the drain current due to substrate bias effects so that equn.(4.4.2) remains valid for more complex drain current expressions than (4.4.1).

in section 4.1, gave a useful drift figure of <1 mV over an 80°C change in ambient temperature. The area of the amplifier was only 24 mils square making it suitable for many CCD peripheral functions.

4.6 Integrated Error Effects

Indications have been given so far, in this section, of the major contributions towards inaccuracy in this particular realisation of the convolution sum. The effects of these error sources at the output of the convolver are now discussed with reference to other work (MacLennan, D.J. et. al., 1976). Unfortunately, unless precise information about both the signal and reference waveforms is available quantitative analysis is not feasible; it is, however, useful to review the general effects of these errors so that potential problem areas may be analysed more fully.

Charge transfer inefficiency (ϵ) in the CCD register has already been shown to degrade signal information at high frequencies; this being a cumulative effect with increasing Nϵ products. In terms of frequency filtering applications, however, the net effect at low frequencies is a shift in the transition edges between pass- and stop-bands of a factor $(1 - \epsilon)$, regardless of the number of filter stages (Buss, D.D. et. al., 1975). The effect upon correlator performance is somewhat more signal dependent but charge transfer inefficiency is known to cause a slump in the correlation peak and a corresponding increase in sidelobe significance. It has been suggested (Buss, D.D. et. al., 1973) that in certain correlation applications an Nϵ product of 2 is tolerable.

Random gain errors in the multiplication process are analogous to tap weight errors in split-gate filters and as such may be expected to impair stop-band suppression in frequency filtering applications. Correlation applications are more tolerant to these errors which become attenuated by the 'processing gain' ($N^{\frac{1}{2}}$) of the filter.

Quiescent offset errors in both the signal and reference channels appear as such at the convolver output for suitably large time-bandwidth products. Taking for example quiescent reference errors, Δr, the net output error will be, $\overline{s.\Delta r}$, where the bar denotes an average value. Clearly, such errors may be reduced or eliminated where either sequence can be chosen to have zero mean value.

The signal distortion term imposed by the finite output admittance of the multiplier driving stages unfortunately precludes a general linear analysis of the resultant error. The effect of this distortion may however be demonstrated in frequency filtering applications where a pure sinusoid is applied at the input. Then, from equn.(4.4.4), we may expect at the output d.c. (zero Hz) and second harmonic distortion components of magnitude $\frac{1}{2}\beta_m|H(o)|/g_o$ and $\frac{1}{2}\beta_m|H(2\omega)|/g_o$ respectively. In correlator applications additional correlation peaks are to be expected whenever a match is detected between the generated harmonic signal component and the reference waveform.

The refreshing and subsequent decay of reference values results in a modulation of the output waveform. As this process is not generally synchronised with the incoming data sequences it appears as a form of noise at the output. Naturally, the magnitude of this noise is directly related to the decay rate and update frequency of the reference samples; these factors determine the maximum number of reference points which may be updated sequentially. Thereafter reference refreshing must take place in parallel blocks.

Although several potential error sources have been identified here it is clear that not all of them will be applicable to any one application of the device. In general, frequency filtering applications requiring large stop-band attenuations, are more sensitive to these errors. Matched filtering (correlation) applications, however are considerably more tolerant to random errors which become attenuated by the 'processing gain' of the filter; generally, the only significant errors are those which correlate with either the signal or reference sequences.

4.7 Implementation

The parallel system circuit principles outlined in this paper have been applied to the design of a monolithic programmable transversal filter for use in a variety of signal processing systems. A block diagram of the proposed device structure is shown in Fig.9, which may be compared directly with the system diagram of Fig.1(b). The tapped CCD delay line feeds one port of a register of MOST multiplier elements. The reference port is similarly driven from the static register the elements of which are updated individually from a multiplexed analogue reference input busbar. Provision is made for the reference values to be feedback linearised at the multiplier ports via a similarly multiplexed feedback busbar.

The feasibility of this filter structure has been verified by construction of a 32-point, hybrid correlator (Mavor, J. et. al., 1977) using a discrete CCD tapped delay line and monolithic MOST multipliers. The output of the device when used as a matched filter to correlate a linear F.M. chirp is shown in Fig.10. This output waveform agrees well with the expected sinc x response and demonstrates the suitability of the device for correlation applications. Frequency filtering is also feasible, although these applications depend greatly upon the accuracy and uniformity of the multiplier elements. Such results will be reported more fully in the future.

The design of a completely monolithic 64-point, programmable transversal filter, shown in Fig.11 has been commissioned and is now in evaluation. A single correlator cell including reference address decode logic is 120 mils long with a width (pitch) of 2.24 mils. The complete integrated circuit measures 180 mils x 130 mils, including an experimental on-chip MOS clock generator and operational amplifier. An updated design is planned on a two-level, polysilicon process with a potentially higher packing density which is expected to yield a 256-point correlator on a chip less than 200 mils square. These devices are cascadable and as such a single board correlator of up to 2000 points may be envisaged.

5. CONCLUSIONS

It is clear from the ever advancing progress of integrated circuit development that advances in the hardware implementation of systems will lead to more compact and effective realisations. In this paper, the application of microcircuits to convolver modules has benefitted from the availability of digital L.S.I. and the potential of custom CCD/MOS technology in analogue form. Although developments according to the latter approach look of great potential presently, undoubtedly, the availability of compact, high performance multiplier hardware would seriously affect the viability of the analogue approach. However, where space and weight is at a premium, the realisation of a convolver with modest performance may still be a useful proposition. This could be required, for example, in mobile surveillance applications where space restrictions limit the hardware to a three chip system: one chip could be a CCD imager, another a CCD/MOS convolver, and the third a post processor (in bipolar or MOS technology). Clearly, the future of the analogue approach is dependent upon the market requirements and the interest shown in compact processors.

This study has indicated that analogue device design has many possibilities when directed to process time-sampled data. It has further shown that an analysis of the errors in such a system is imperative before the operation of an analogue convolver can be properly assessed. Such work needs to be extended in the light of other prototype measurements before the full limitations of the approach can be determined.

Where the precision of processing is of paramount important, LSI-based digital hardware can be used to effect to produce a relatively compact convolver. Single-chip multipliers and mass memory are increasingly making the digital approach all pervading for virtually any application. The advantages of the prototype Deltic processor described in this paper were clearly evident, and had the merit that all operating parameters could be extended very easily. Thus, in the longer term, the digital approach may supersede analogue approaches, but, analogue device research will continue to show interesting possibilities.

6. ACKNOWLEDGEMENTS

7. REFERENCES

BOYLE, W.S. and SMITH, G.E., 1970, "Charge-coupled semiconductor devices", BSTJ, 49, pp.587-593.

BOSSHART, P., 1976, "An integrated analogue correlator using charge-coupled devices", IEEE ISSCC 76, pp.198-199.

BUSS, D.D., COLLINS, D.R., BAILEY, W.H. and REEVES, C.R., 1973, "Transversal filtering using charge transfer devices", IEEE J. of Solid-State Circuits, SC-8, pp.138-146.

BUSS, D.D., BAILEY, W.H., HOLMES, J.D. and HITE, L.R., 1975, "Charge transfer device transversal filters for communication systems", Microelectronics Journal, 7, pp.46-53.

COPELAND, M.A., 1977, "Interaction between microprocessors and custom LSI," AGARD Lecture Series No.87, "Microprocessors and their applications", New York.

DENYER, P.B. and MAVOR, J., 1977, "Design of CCD delay lines with floating-gate taps", IEE J. on Solid-State and Electron Devices, 4, (to be published).

HARP, J.G., VANSTONE, G.F., MACLENNAN, D.J. and MAVOR, J., 1975, "Analogue correlators using charge-coupled devices", Proc. 1975 CCD Applications Conference, San Diego, pp.229-235.

HERRMANN, E.P., GANDOLFO, D.A., BOORNARD, A. and STEPPS, D.B., 1976, "CCD programmable correlator", Proc. 1976 Int. Conference on the Tech. and Applications of CCD's, Edinburgh, pp.232-237.

MACLENNAN, D.J. and MAVOR, J., 1975, "Novel technique for the linearisation of charge-coupled devices", Electronics Letters, 11, pp.222-223.

MACLENNAN, D.J., MAVOR, J. and YEOW, Y.T., 1976, "Errors in programmable CCD transversal filters and correlators", Proc. 1976 Int. Conference on the Tech. and Applications of CCD's, Edinburgh, pp.259-268.

MAVOR, J., JACK, M.A., SAXTON, D. and GRANT, P.M., 1977, "Design and performance of a programmable real-time charge-coupled device recirculating delay-line correlator", IEE J. on Electronic Circuits and Systems, 4, (to be published).

MAVOR, J., ARTHUR, J.W. and DENYER, P.B., 1977, "Analogue CCD correlator using monolithic MOST multipliers", Electronics Letters, 13, pp.373-374.

SQUIRE, W.D., WHITEHOUSE, H.J. and ALSUP, J.M., 1969, "Linear signal processing and ultrasonic transversal filters", IEEE Trans. MTT-17, pp.1020-1040.

TASCH, A.F., BRODERSEN, R.W., BUSS, D.D. and BATE, R.T., 1973, "Dark current and storage-time considerations in charge-coupled devices", Proc. 1973 CCD Applications Conference, San Diego, pp.179-188.

WESTE, N. and MAVOR, J., 1977, "MOST amplifiers for performing peripheral integrated circuit functions", IEE J. on Electronic Circuits and Systems, 5, (to be published).
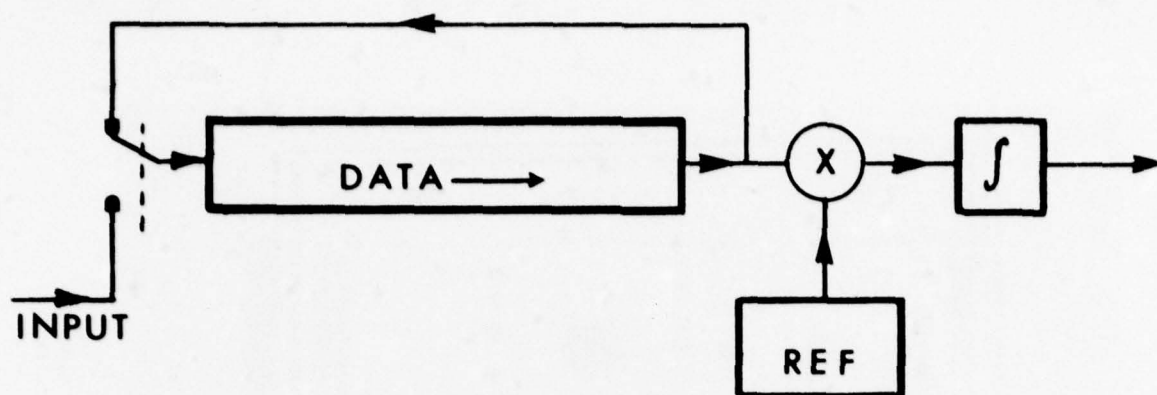
(a)
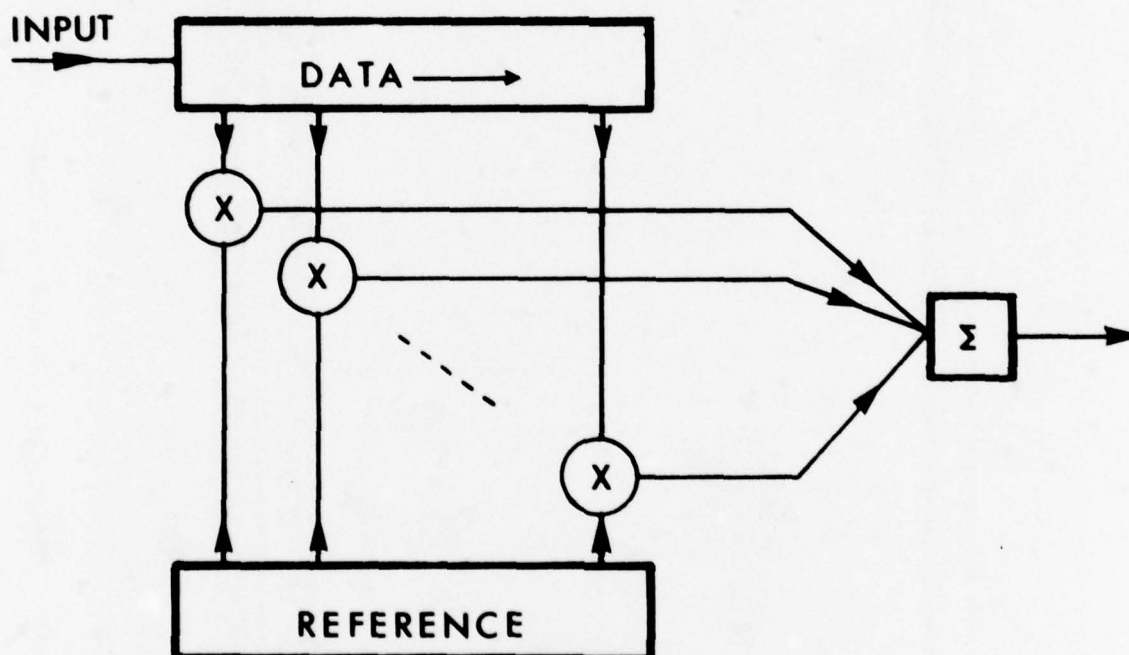
(b)

FIG.1.   Serial and Parallel forms of implementing the convolution sum.
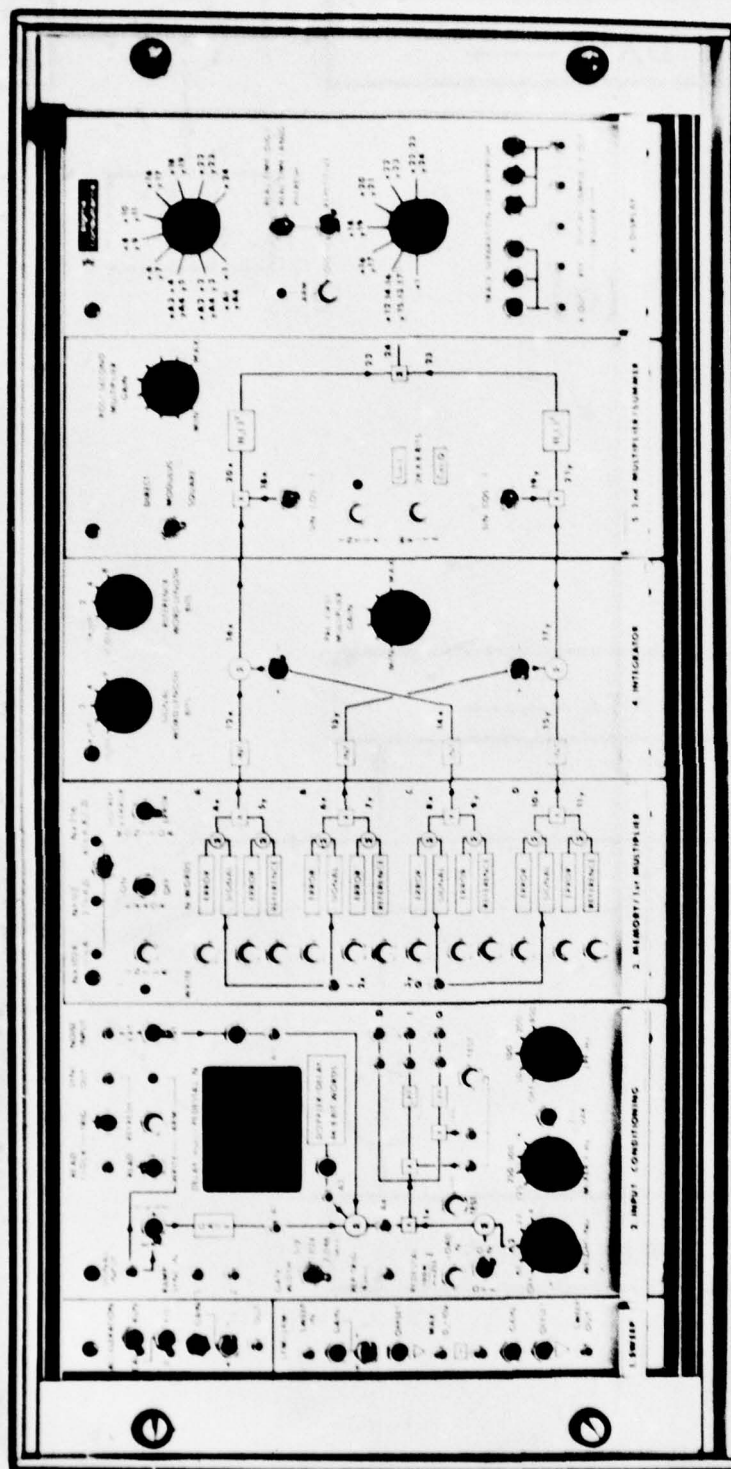
FIG.2. Photograph of a 1000 point serial convolver based on the DELTIC principle. (Courtesy Sigma Consultants).
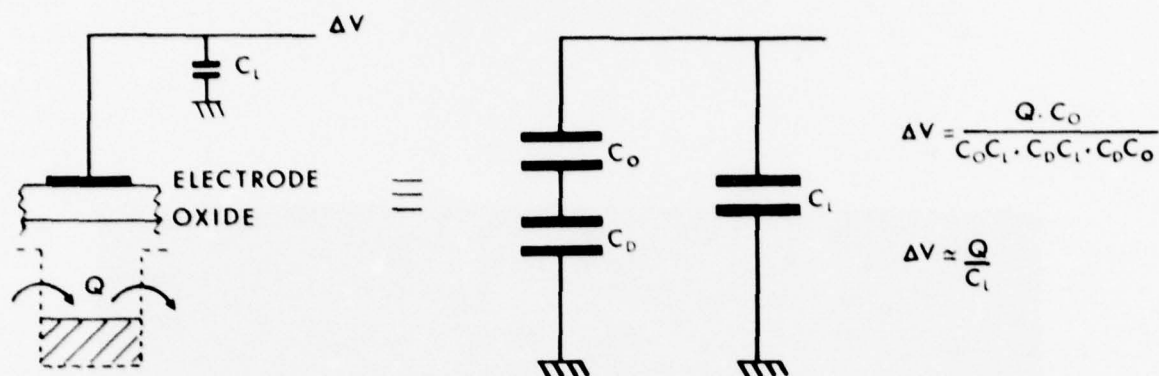
FIG. 3.   Principle of Floating Gate Reset charge sensing scheme.

$$\Delta V = \frac{Q \cdot C_O}{C_O C_L \cdot C_D C_L \cdot C_D C_O}$$

$$\Delta V \simeq \frac{Q}{C_L}$$
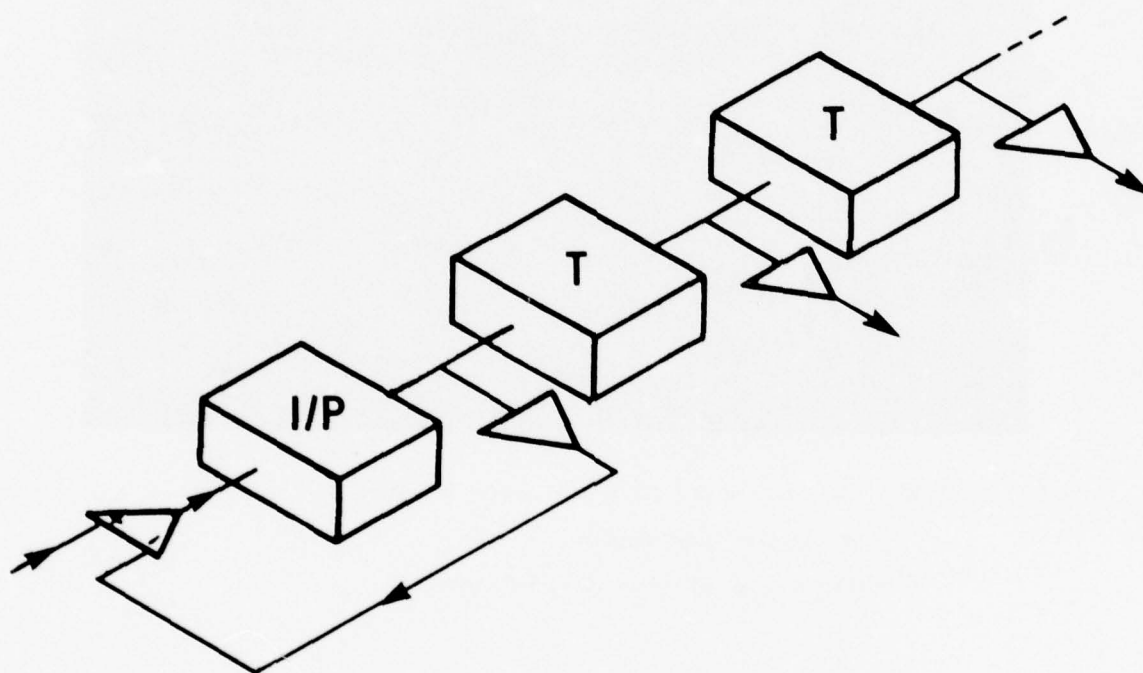
FIG. 4.   Feedback Linearisation of the CCD transfer function.
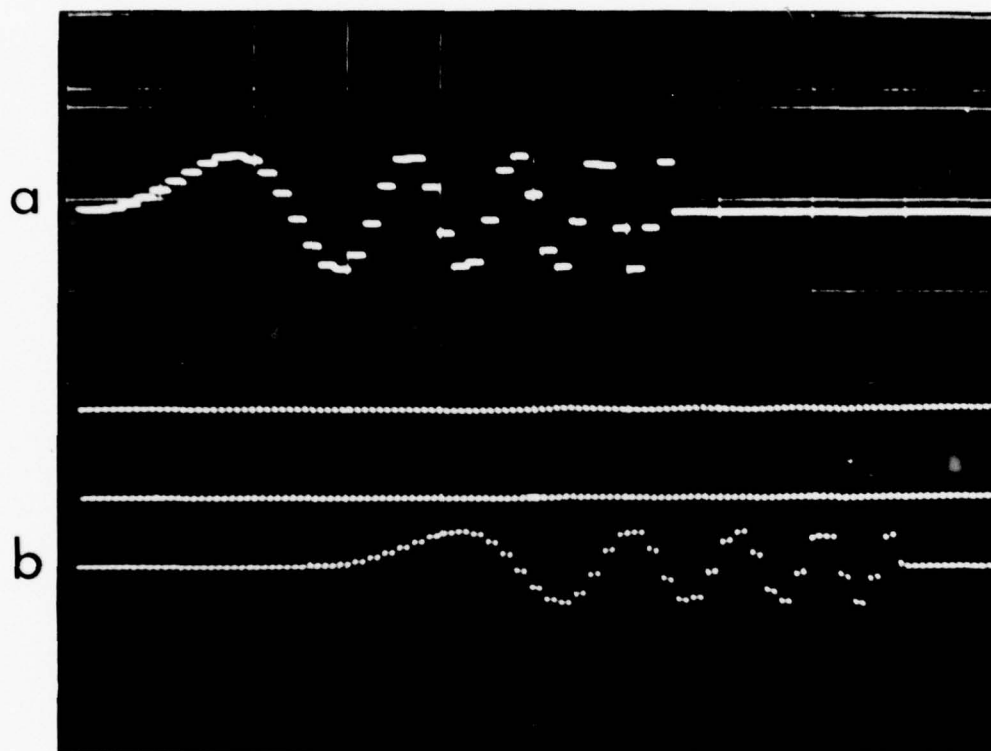
FIG.5. Photograph of FGR tap output showing a delayed chirp waveform.

(a) Sampled Input Waveform.

(b) Delayed Tap Output (2 V, 200 μsec).
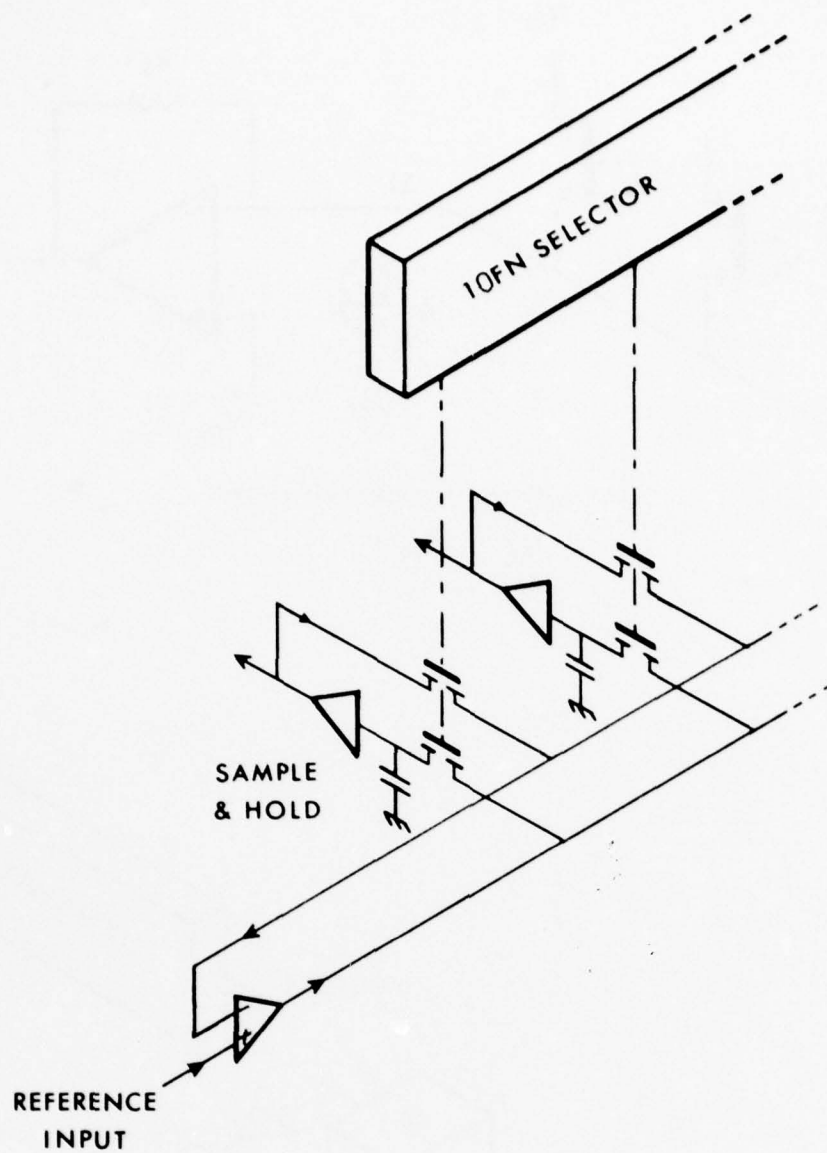
FIG.6.   Digitally multiplexed capacitive analogue reference register with feedback linearisation.
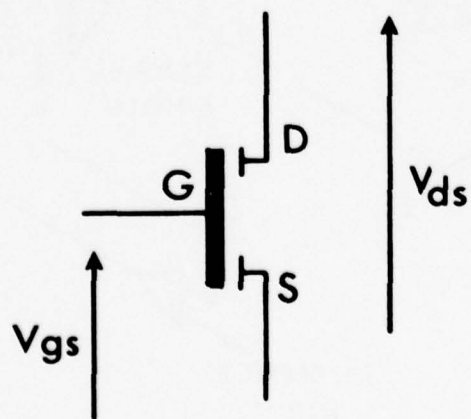
FIG.7.   MOS transistor nomenclature.
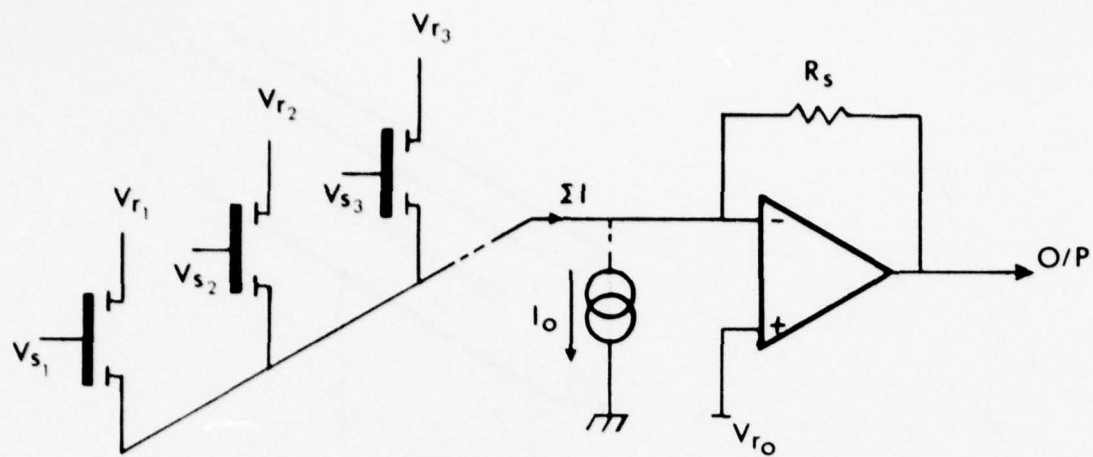
FIG.8.   Multiplier summing arrangement.

$$V_S = v_s + V_{s_0}$$
$$V_r = v_r + V_{r_0}$$
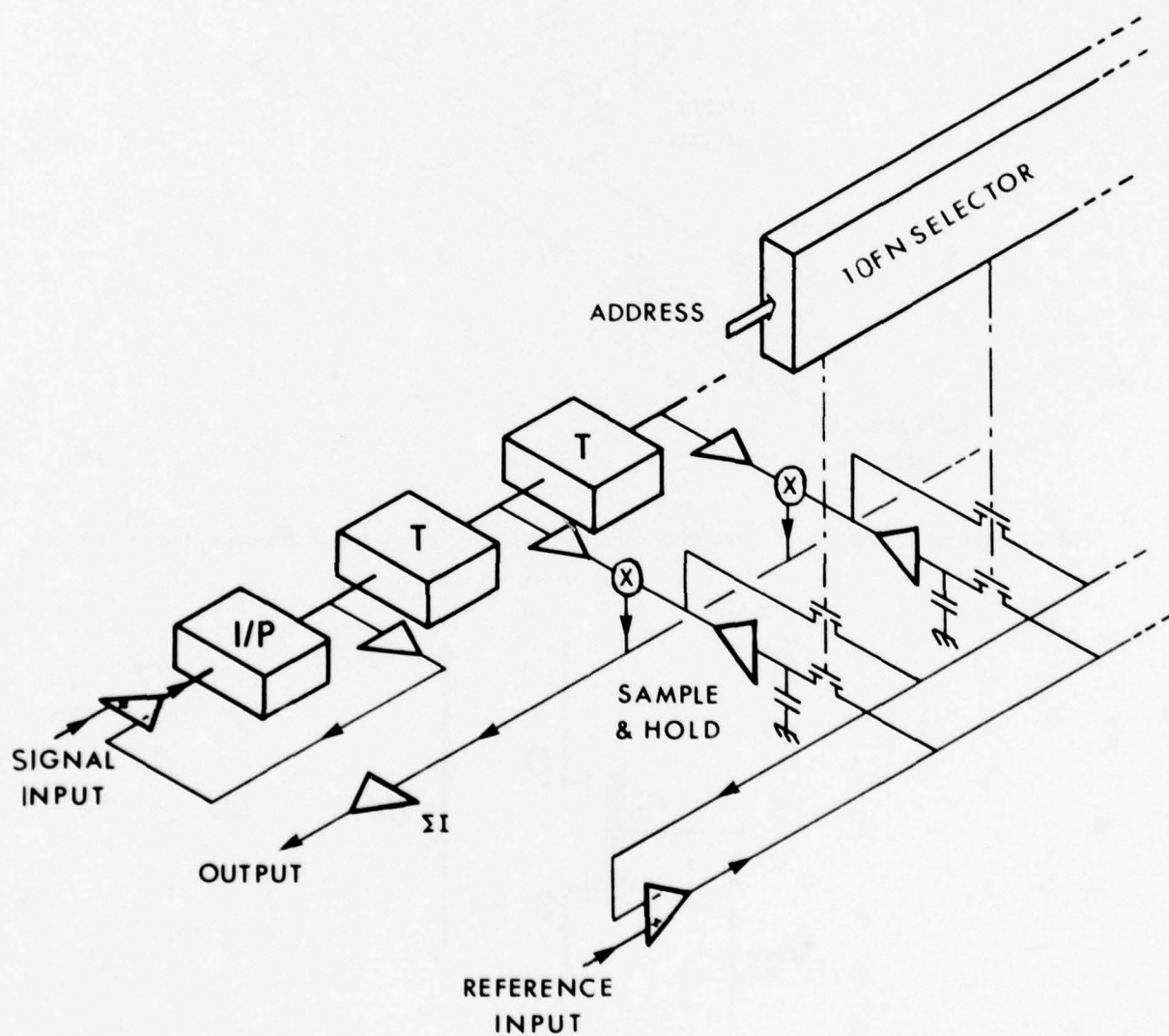
FIG.9.   Block diagram of a programmable transversal filter.
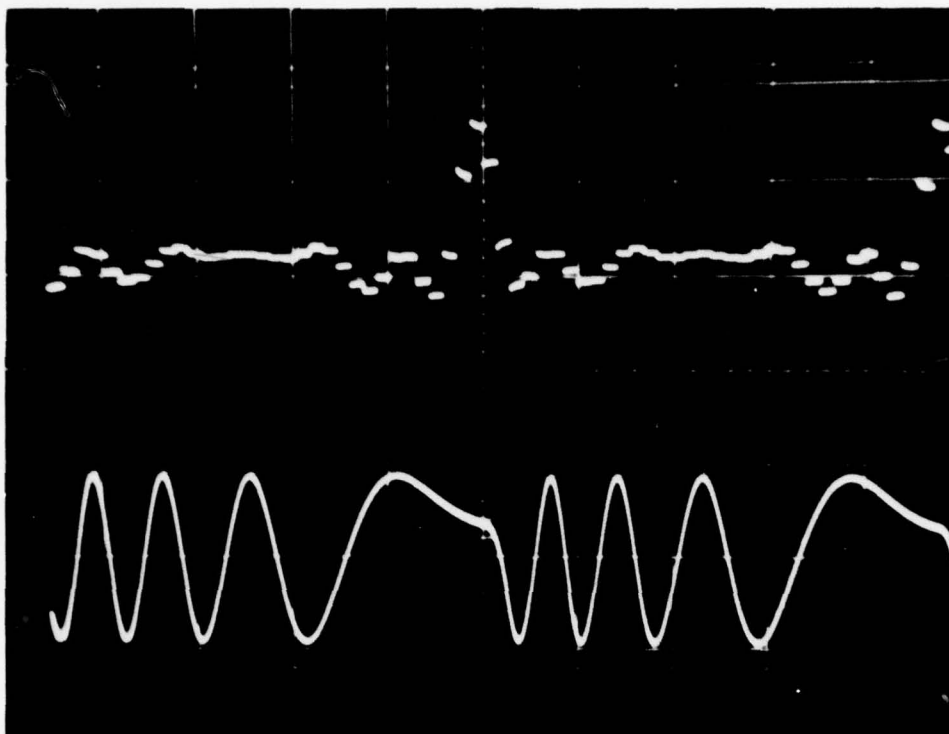
FIG.10. Matched filtering of chirp waveform using 32 point hybrid convolver.
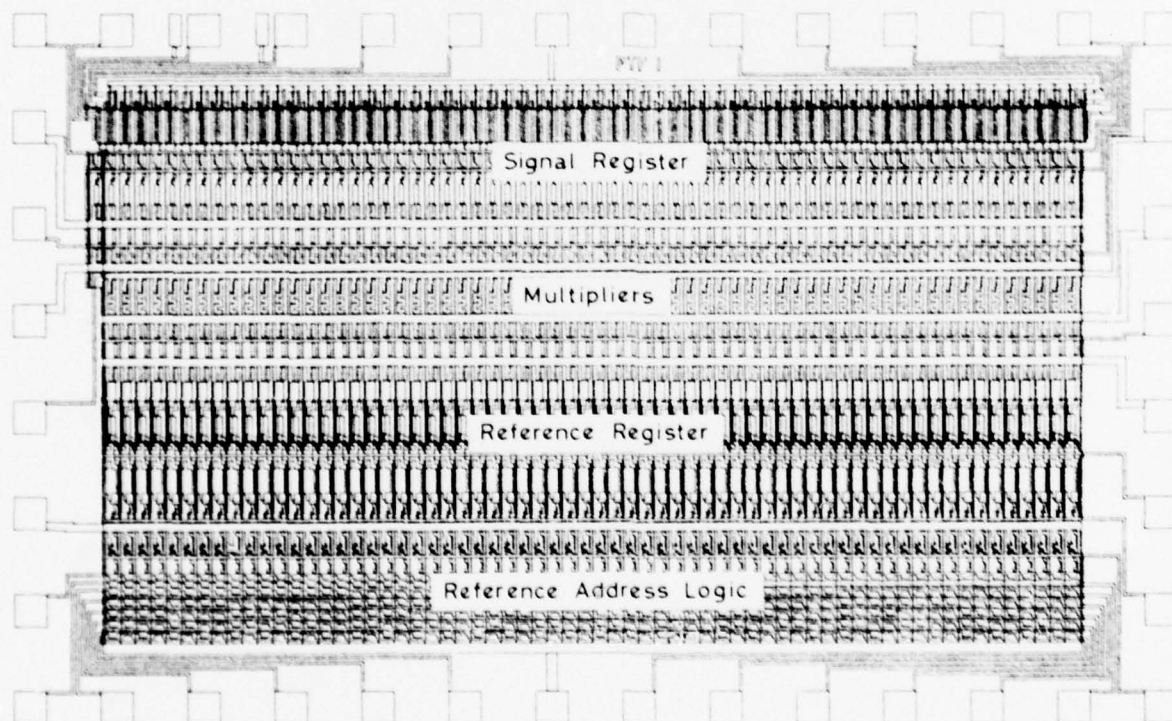
FIG.11. Drawing of a monolithic 64 point analogue-analogue programmable transversal filter.

DISCUSSION

Copeland

Can you comment on your dynamic range. What is the signal-to-noise ratio? Is there any consideration of linearity in your dynamic range for the moment?

Author's Reply

With feedback linearization, we find that the linearity is preserved in the dynamic range.

The dynamic range was measured through the autocorrelation of square wave, providing a peak signal, and we then detect the effective RMS noise level.

The limitation at the present is probably the decay on the sample hold capacitor. This represents the major problem. The breakthrough between the driven clock gate and the flooding gate on the CCD probably reduces the dynamic range and causes a certain amount of spurious signals.

Tournois

If you rely on using your device in adaptive signal processing (for example, sidelobe cancellations) you need a precision on your weightings coefficients. Which precision can you obtain in your multiplier?

Author's Reply

Difficult to quantify. We have looked at the autocorrelation of two square waves and the triangular output is nearly perfect, which means within one percent.

We need 500 samples to measure before making a judgement.

Name unknown

Can you comment on power consumption?

Author's Reply

The total power is 350 milliwatts.

A HYBRID SAW/CCD SIGNAL PROCESSOR

William H. Conway
Rockwell International
Electronics Research Center
3370 Miraloma Avenue, Anaheim, CA 92803

Dirk R. Klose
US Army Electronic Warfare Laboratory (ECOM)
DRSEL-WL-S
Fort Monmouth, New Jersey 07703

ABSTRACT

This paper discusses the combining of SAW and CCD technologies to produce a hybrid signal processor capable of performing bandwidth compression on signals of short time duration while maintaining their phase relationships. The devices used in fabricating a feasibility brassboard are described and test results influencing the design of such hybrid processors are presented.

1. INTRODUCTION

This paper presents the results of a program to develop a "SAW/CCD High Speed Signal Digitizer" for the U.S. Army Electronics Command by the Electronics Research Division of Rockwell International Corporation. The following paragraphs will describe the particular device used to develop an exploratory model of this advanced signal processor. The indicated processor represents a unique marriage of SAW and CCD capabilities to accomplish a signal processing function.

2. TECHNICAL BACKGROUND

Surface acoustic waves and charge coupled devices are two relatively recent technologies that have matured at about the same time. When applied to the area of signal processing, they have frequently been viewed as competitors. In fact, there are several areas of signal processing where they are complementary--together extending the processing power to a point that one alone could not achieve.

Their ability to complement arises directly from their characteristics. SAW devices are passive, passband, fixed coded, large dynamic range components. They, in general, have wide applications to analog signal processing functions of large time-bandwidth products and specific fixed operations. SAW technology achieves large time bandwidth products via short interactions time processing of wide bandwidth signals. The passive nature of SAW devices gives them a large linear dynamic range capability. Their fixed coded structure has resulted in good parameter reproducibility. SAW devices operate directly at IF/RF frequencies resulting in significant reduction of equipment size and complexity.

CCDs are active, baseband, variable rate, fixed dynamic range components. They are sampled analog processing devices featuring long interaction time processing at restricted bandwidths. Their linear dynamic range is a function of the speed at both extremes of operation. The high speed signal is limited by the slow rate of the devices, and the low speed is limited by thermal filling of the wells.

Individually, the characteristics of SAW and CCD devices are contradictory, yet many systems have requirements which contain just such built-in internal contradictions.

This paper discusses the combining of SAW and CCD technologies to produce a hybrid system capable of performing bandwidth compression on signals of short time duration while maintaining their phase relationships. A feasibility brass-board system has been built using proprietary Rockwell International SAW and high speed, bulk channel CCD components. All other devices used consisted of low cost off-the-shelf components. The program objective was to develop a state-of-the-art processor with currently available hardware.

3. HYBRID SAW/CCD SIGNAL PROCESSOR

The processor discussed in this paper is a pulse burst analog to digital converter with a high speed serial input and a low speed serial output. The processor will accept a 5 usec pulse burst from an IF amplifier with 15 MHz bandwidth at a 30 MHz center frequency.

The IF input will be stored and time multiplexed by a surface acoustic wave tapped delay line as shown in Figure 1. The SAW line converts the serial analog input to four parallel synchronous analog outputs. Alternately, it can be viewed as providing four data blocks 1.25 usec long (the time delay between consecutive taps). In addition, the SAW line provides band limiting as an IF filter complementary to the IF input.

Each data block is fed into a 128 Register CCD sampled analog delay line through a buffer amplifier. The buffer stage provides a signal level which matches the input impedance and signal level requirements of the CCD in order to minimize any overall loss of dynamic range. Since the highest frequency component of the input signal is 37.5 MHz, the waveform can be preserved if it is sampled at greater than 2.5 x 37.5 MHz = 93.75 MHz. For this reason, each CCD register is driven by a 102.4 MHz input clock. Since each CCD is sampling a 1.25 usec block at the 102.4 MHz rate, each CCD will store 128 analog samples for a total of 512 signal samples.

These particular frequencies have been chosen to maintain a binary number relationship between the input sample periods and the sampling readout periods. (Considering 5 us to be 1/2 of 10 us and 1.25 us to be 1/8 of 10 us, they and the 102.4 MHz and 128 samples all share the power of 2 relationship). This greatly simplifies the associated control logic and could even allow the clocks to be phase locked to a 100 KHz frequency standard.

When the entire data block is loaded into the CCDs, the clock shifts to low speed for the readout phase. The output of the CCDs is scanned by a four input multiplexer and presented to the sample and hold module and then to the analog to digital converter. The time relationship of the original signals are now multiplexed so that the four input channels appear as they have been stored in the CCDs, (i.e., samples 1, 129, 257, 385, 2, 130, 258, 386,...). These signals are then loaded into a digital memory in a manner that demultiplexes them so that adjacent sample period appear in sequential memory locations. This memory can be sequentially read out in two ways: (1) through a digital-to-analog converter to provide an analog representation of the input waveform, or (2) through an 8-bit digital port. It can also be randomly accessed through an external address bus.

4.   DEVELOPMENT APPROACH

The approach used in this program was to use previously developed components whenever possible. These components have been assembled to form an end item that has capabilities far greater than any of the individual items. The drives and control logic consists of standard ECL logic integrated circuits. The operating speeds of this processor required the high speeds obtainable with ECL for these functions. The low speed read out multiplexer and A-D converter consist of standard hybrid bipolar modules. The memory is composed of standard MOS static memory elements. The custom components used for this processor, are the SAW tapped delay line and the CCD sampled analog delay line.

The tapped delay line shown in Figure 2 used in the SAW/CCD hybrid processor consists of nine half-phase reversed (1/2 PRT), transducers with eight fingers each. These transducers are separated by .4418 cm., giving a delay time on lithium niobate (LiNbO$_3$) of 1.25 usec. LiNbO$_3$ was chosen as the piezoelectric substrate material in order to minimize insertion loss at the large fractional bandwidth required.

Bandpass data obtained is shown in Figures 3 and 4. The passband shown in Figure 3 was obtained with untuned transducers while the data shown in Figure 4 was taken with both transducers tuned to center frequency (30 MHz). It is readily apparent that tuning reduces the insertion loss by about 10 dB, and both narrows and flattens the passband. The flat passband between the three dB points of the tuned device is ~11.5 MHz, while the corresponding untuned passband is ~15 MHz. The tuned device is clearly preferred however, because of low insertion loss (~20 dB), and flat passband. (The untuned case has a dip in the center about 5 dB below maximum). Both tuned and untuned cases agree closely with theoretical calculations based on the "cross-field" transducer model. Some ripple will be noted in the passband. This is due primarily to bulk and other reflections. As no special effort was made to reduce these effects during the initial investigations, the presence of the ripple is not surprising. In the final assembly, care was taken to minimize these spurious effects.

In order to meet the high frequency requirements of this signal processing task, a bulk channel (peristaltic) CCD was chosen. Such devices have been fabricated and operated up to frequencies of 230 MHz. In the 100 MHz range, charge transfer efficiencies in excess of 0.99994 per gate have been observed with these devices.

The PCCD was first developed at Rockwell for a task very similar in operation to the one being addressed in this paper: The capture of information from a high frequency burst. This necessitates operating the CCD in a high frequency mode during information read-in and then switching to a low frequency mode for information read-out. Due to the success of this earlier undertaking, the existing high-frequency CCD units were chosen to minimize the cost to this program and to allow attention to be focused on the operation of the SAW/CCD hybrid processor.

Normally, CCD devices transfer the final portion of charge under a given gate from a location which is so near to the transfer gate surface that the effect of the transverse field is negligibly small. This effect limits the operating speed with which good transfer efficiency can be obtained to around 20 MHz. In order to operate at higher frequencies, Esser introduced the peristaltic CCD which transfers charge in a channel several microns below the surface. This configuration was obtained by using epitaxial layers grown on silicon substrates. Because the transverse electric field is very strong at the charge transfer area, the charge can be moved very rapidly out of each cell and, therefore, this type of CCD has been operated at very high frequencies.

The device used here is a 128-cell, four-phase CCD. It was fabricated using an N-type epitaxial layer grown on P-type silicon. The dopings of these layers were 1×10^{15} and 5×10^{14} cm^{-3}, respectively. A cross sectional view of the design is shown in Figure 5. A four-phase overlapped A propagation gate structure was employed as the driving structure. Because the device was intended for high speed operation, gate length (in the direction of propagation) was made as small practicable, i.e., 0.2 mils. The total cell length consisting of 4 gates is 0.8 mils. The channel width is 10 mils in order to provide sufficient charge per cell. Referring to the figure, the first gate on the left is used as an input control electrode. A built-in FET amplifier has been designed into the structure as an output. The complete chip consisting of two PCCDs and auxiliary test circuitry is shown in the photo of Figure 6. The P$^+$ diffusion around the CCD channel acts as a channel stop in order to contain the active charges. Outside the P$^+$ channel stop is an N$^+$ diffusion which functions as a leakage sink.

5.   CONFIGURATION AND PERFORMANCE CONSIDERATIONS

The data readout portion of the processor shown in Figure 1 can be configured in several different ways. The simplest conceptually is single channel end-to-end readout using the CCD as a true analog shift register. The output of each register would be fed to the input of the next register and finally to the A-D converter. Such a configuration automatically restores the time order of the data samples.

A low speed all TTL version of the logic and clock devices for CCD operation was constructed during the early portion of the program to gain working experience with the CCDs operating in the fast load, slow unload mode, and to work out the details of interfacing the SAW and CCD devices. This information was also needed to design the interstage networks that would multiplex between the SAW input in the high speed load mode and the output of the preceding CCD in the low speed readout mode.

During these tests, it was determined that a slightly different operating point (bias voltage setting). is required to optimize each of these two loading conditions. The reason for this can be readily seen for at least one setting by comparing the input currents required during the load mode. The CCD can be considered to require a certain quantity of charge, q, for each sample of the input voltage entered into the "analog delay line". If N samples are to be entered, (in this case 128), the total charge required will be Nq (i.e., 128q). independent of the time required to deliver this charge. If the time period over which this charge is delivered is varied, (in this case by a factor of 128- from a 102.4 MHz input sample rate to a 800 KHz output sample rate), the input current to the device must also be varied. This current is typically supplied by a constant current source to reduce the effect of temperature variations of the input diode. Therefore, the constant current generator would have to be varied with the input clock rate. This variation can be accomplished by coupling a portion of the clock into the input current generator to convert it to a "pulsed constant current" source.

Observations made during the operation of the low speed CCD operation indicate that there are also some interactions between the input current variations and the other bias parameters involved in the CCD that would preclude the simple adjustment of current being the only corrections to be made to compensate for changing the input rate.

Because of these indications, it became desirable to re-evaluate the method of handling the data during the unload phase of operation to see if there was a simple, reliable way to avoid feeding the data from one CCD to another, and still retain the time order of the output digital data.

One method was to clock out one CCD at a time. This retains the time order of the data, but would require that the number of CCDs be doubled since two are packaged to share the same clock drive signals. This would also require additional driver circuits for these extra CCDs.

Another simple alternative method is to clock all the CCDs together and multiplex their outputs to the A-D converter (or provide separate A-D converters for each CCD channel). This is a simple hardware solution, but will result in an output that has the four data channels in a time multiplex format. If the data is to be presented directly to a computer for analysis, this does not present a significant problem. However, if the data is to be examined manually, it provides a small complication.

If the data is to be examined manually, either by stepping through with a digital readout, or to be recorded using an analog device, (i.e., strip chart or oscilloscope). it is further necessary to provide some storage media for the signal since it must be read from the CCD in a matter of a few milliseconds to avoid any problem from thermal filling of the wells. This memory can then be used to demultiplex the data by partitioning the memory space into four sections and loading data from each CCD into the proper section. The memory can be read out by cycling through from one end to the other, thereby restoring the proper time sequence to the data. In addition, since this memory is a conventional digital memory, and not subject to time decay, it can be held for any convenient time and/or cycled continuously to provide a simple repetitive readout of the input signal. This can be passed through a D-A converter and then displayed on a conventional oscilloscope spectrum analyzer or other analog devices.

The last alternative method is a variation of the foregoing where separate parallel channels of A-D converters and memory replace the multiplexing system. The main advantage of such a system is to increase the readout speed of the processor but at an increase in cost.

The multiplex version as shown in Figure 7 has been used in this demonstration project to minimize costs.

6.   PERFORMANCE LIMITATIONS

The limits of performance are determined by three major considerations. The first of these is the charge transfer rate. This is, in practice, limited by the maximum drive rate that can be applied to the transfer gates.

The transfer gate structure of a CCD appears as a relatively large capacitance that must be charged to different voltage levels to accomplish the transfer of charge down the CCD propagating structure. When this capacitance must be charged at a high frequency, the process requires switching large currents and can consume a considerable amount of power. Typically, the drive circuits for a CCD will dissipate up to 20 watts per channel when operating at 100 MHz.

As a result of this power dissipation level, it has been the general practice to only operate the drive circuits during actual device operation periods and keep the circuit in a stand-by mode at other times. This results in a complicated logic being required and can lead to loss of data under some circumstances. Because of this, an attempt was made during this program to design a drive circuit that would develop the necessary drive voltages at 100 MHz and not dissipate large amounts of D.C. power when operating at low drive frequencies. This would eliminate the requirement to start and stop the drivers and would result in lower overall power dissipation in the circuit.

The most promising circuit investigated thus far, is using a complementary emitter follower circuit to drive the load capacitance. This circuit only dissipates power during the actual transition period between high and low states. If the input signal is driven to nearly the supply rails, very little power is consumed during the high or low states. Major effort in this direction has been to identify complementary transistors suitable for this type of operation.

The second performance limitation is due to the charge transfer efficiency. This factor has been mentioned previously to be 0.99994 for the device being used here. When this level of efficiency is applied to all 128 transfers, the overall efficiencey is still 0.9923, which is sufficient for this application.

The third performance limitation is device noise characteristics. These include thermal filling as well as clock feedthrough. Any other spurious signals that may be encountered also add to the noise level.

7.   CONCLUSIONS

In this paper, the feasibility of combining SAW and CCD devices has been discussed, and a hybrid high speed signal digitizer has been characterized. Initial test results confirm feasibility. Further test data will be available at the time of presentation.

The major problem encountered to date, resides in the driver circuits for the propagation gates. Obtaining 15v drive levels at 100 MHz rate requires power video amplifiers capable of 1 amp current surges. The results of this effort indicates that a significant advantage is obtained by combining two complementary technologies into a hybrid form to achieve both wide bandwidths and large interaction time processing. While a relatively simple signal capture and digitization function was considered in this paper, the real potential advantage of this hybrid technology concept lies in its application to complex signal processing functions requiring both wide pre-detection bandwidths and long signal interaction times.
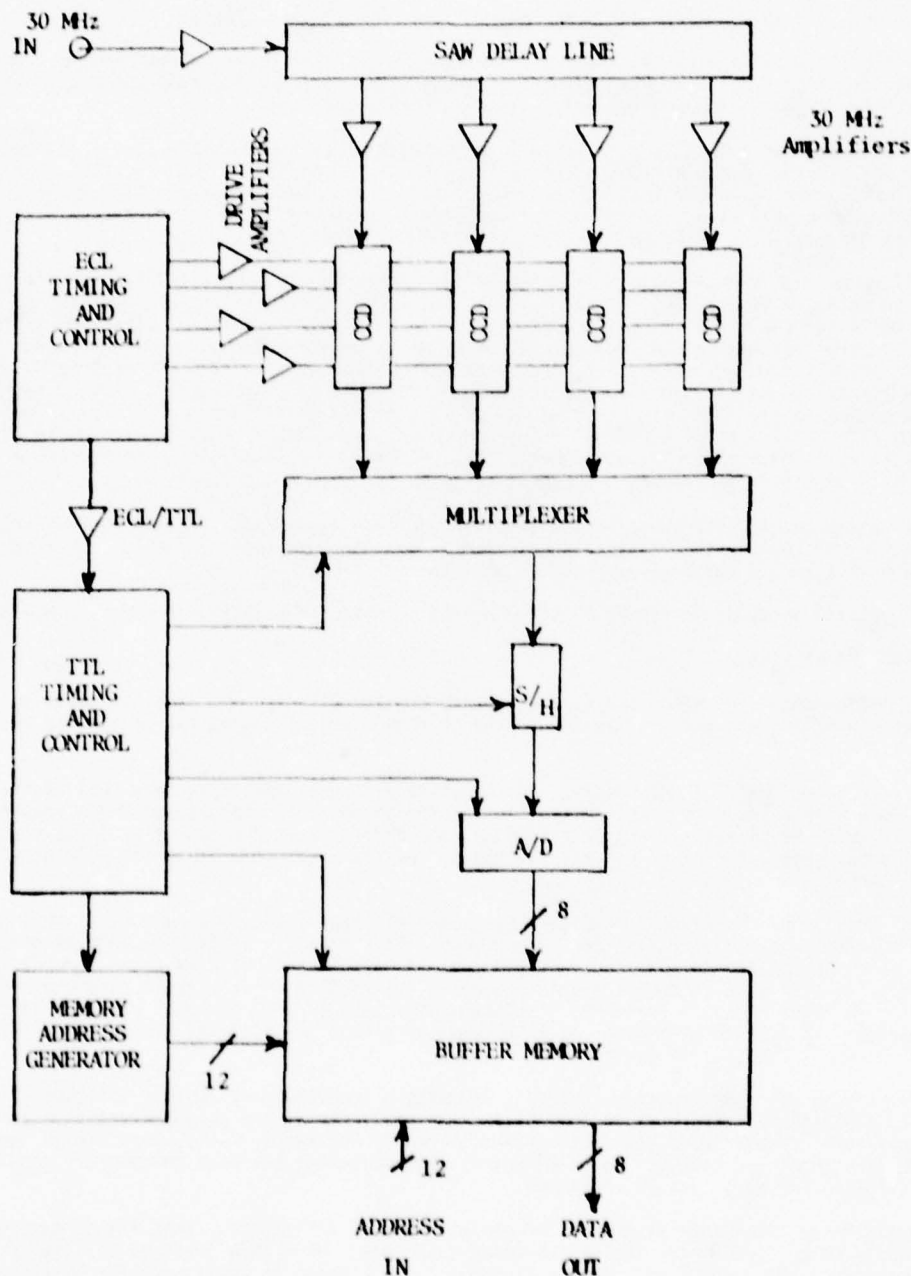
Fig 1   SAW/CCD hybrid processor block diagram
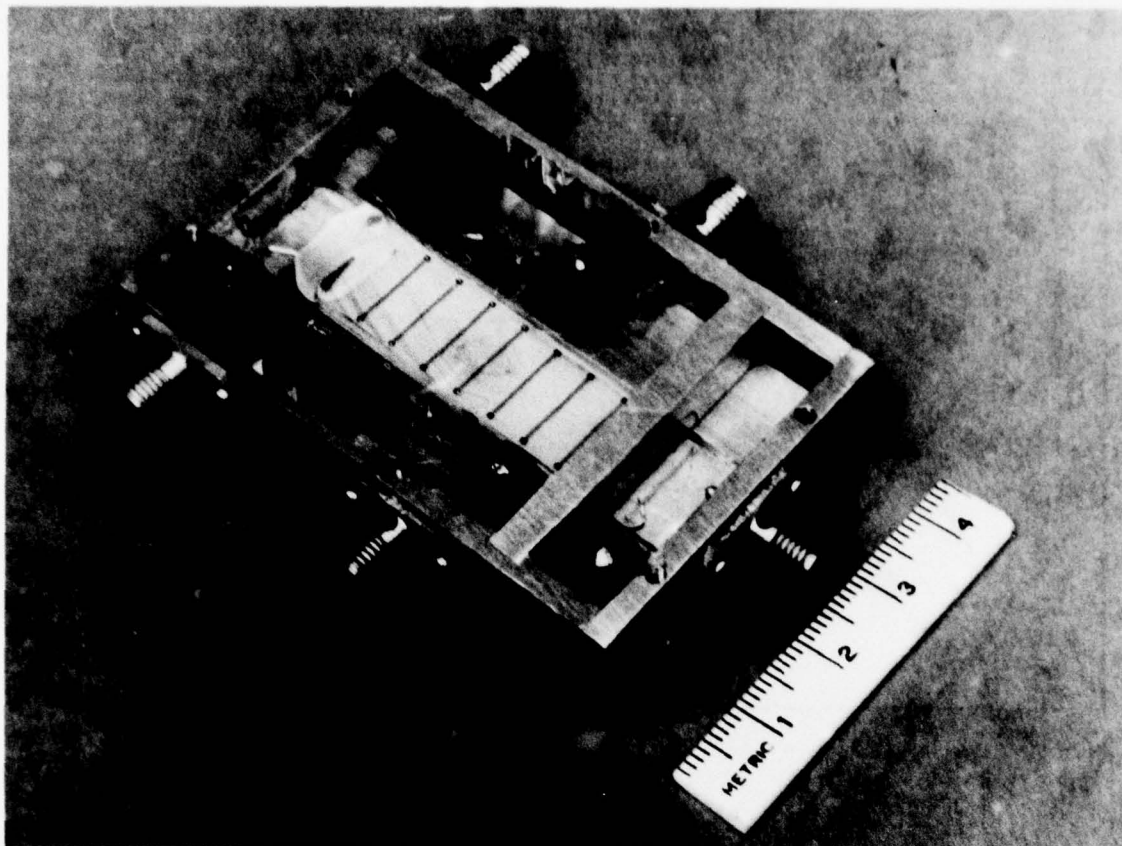
Fig.2   SAW tapped delay line
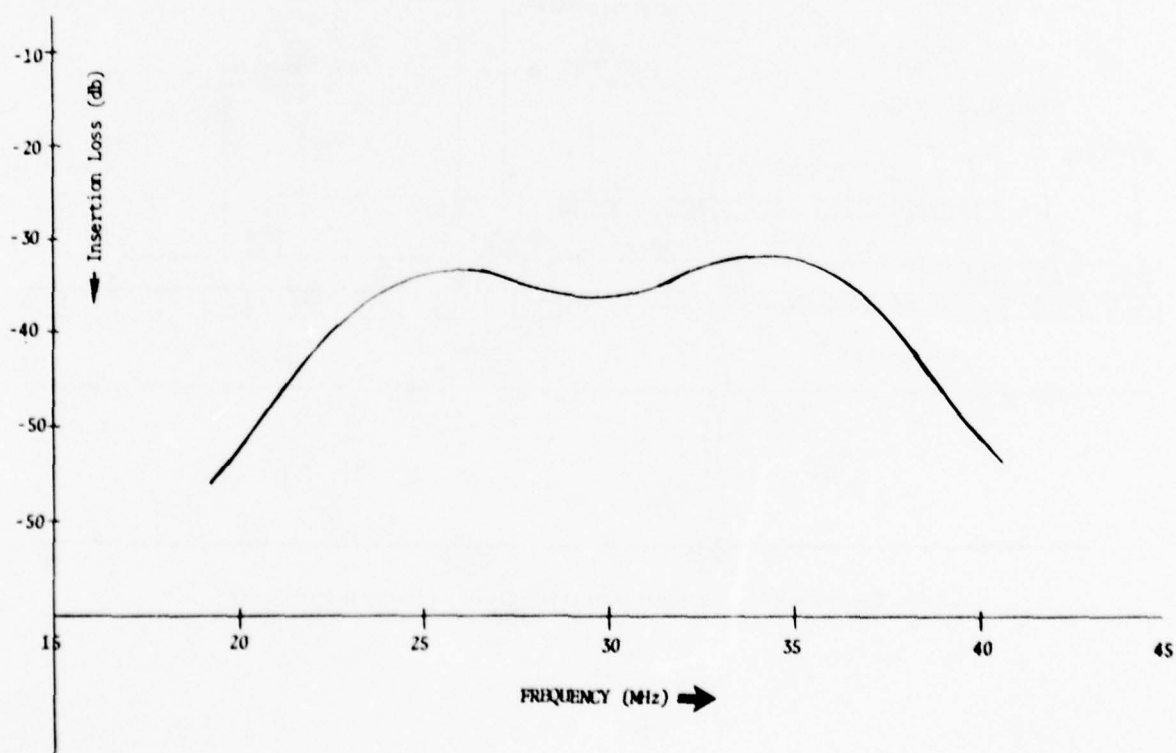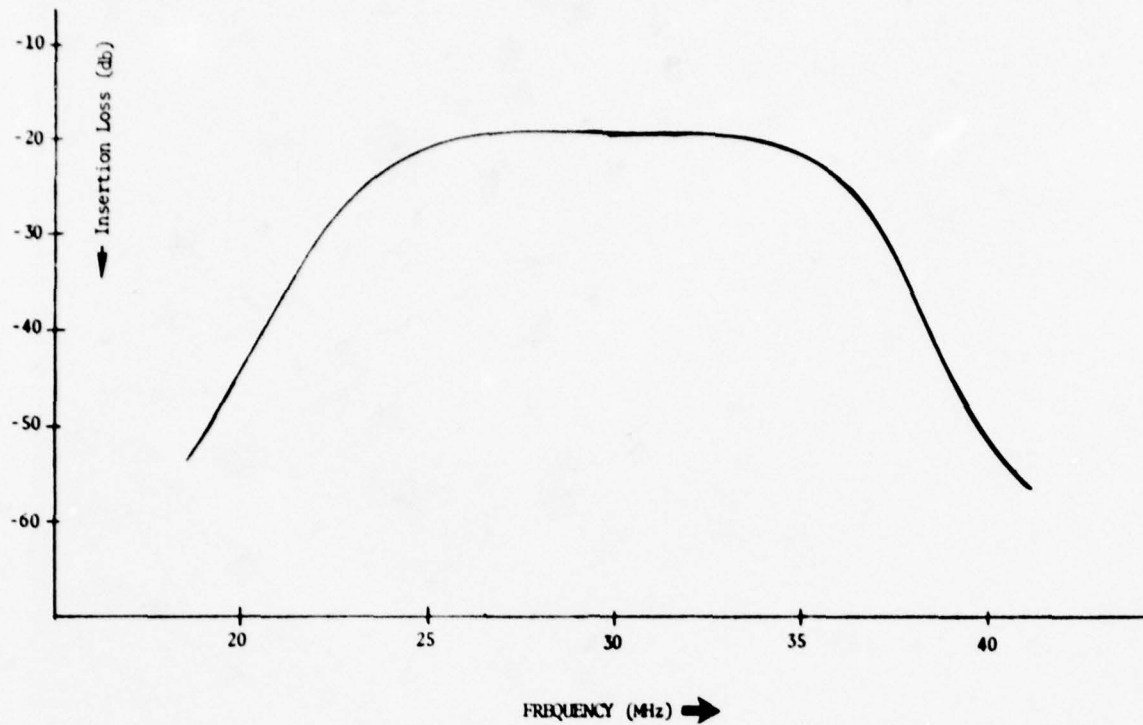

Fig.3   Untuned passband
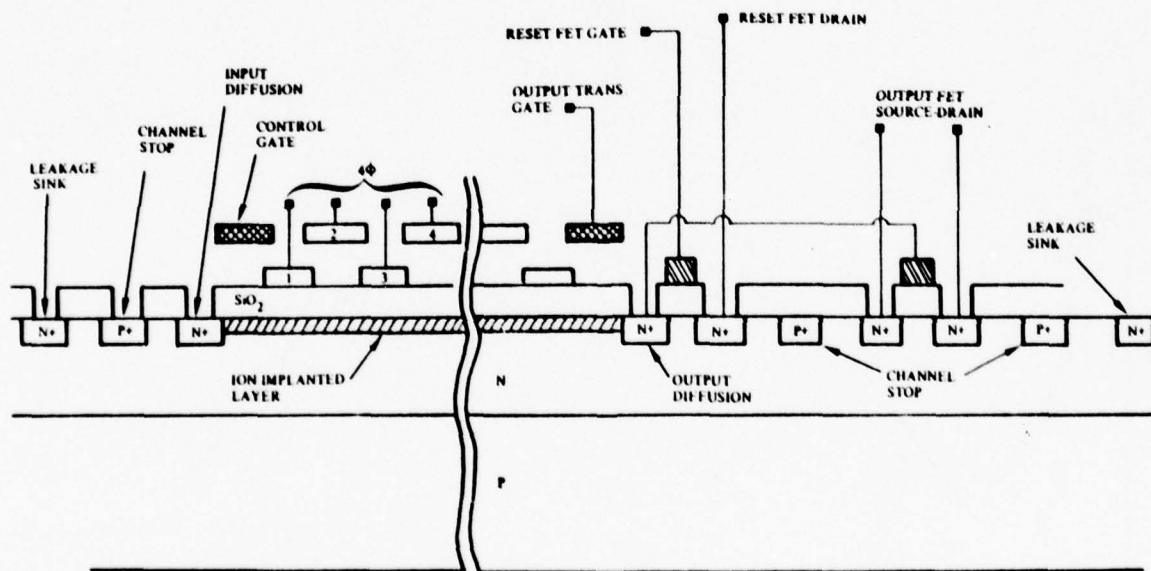
Fig.4 Tuned passband

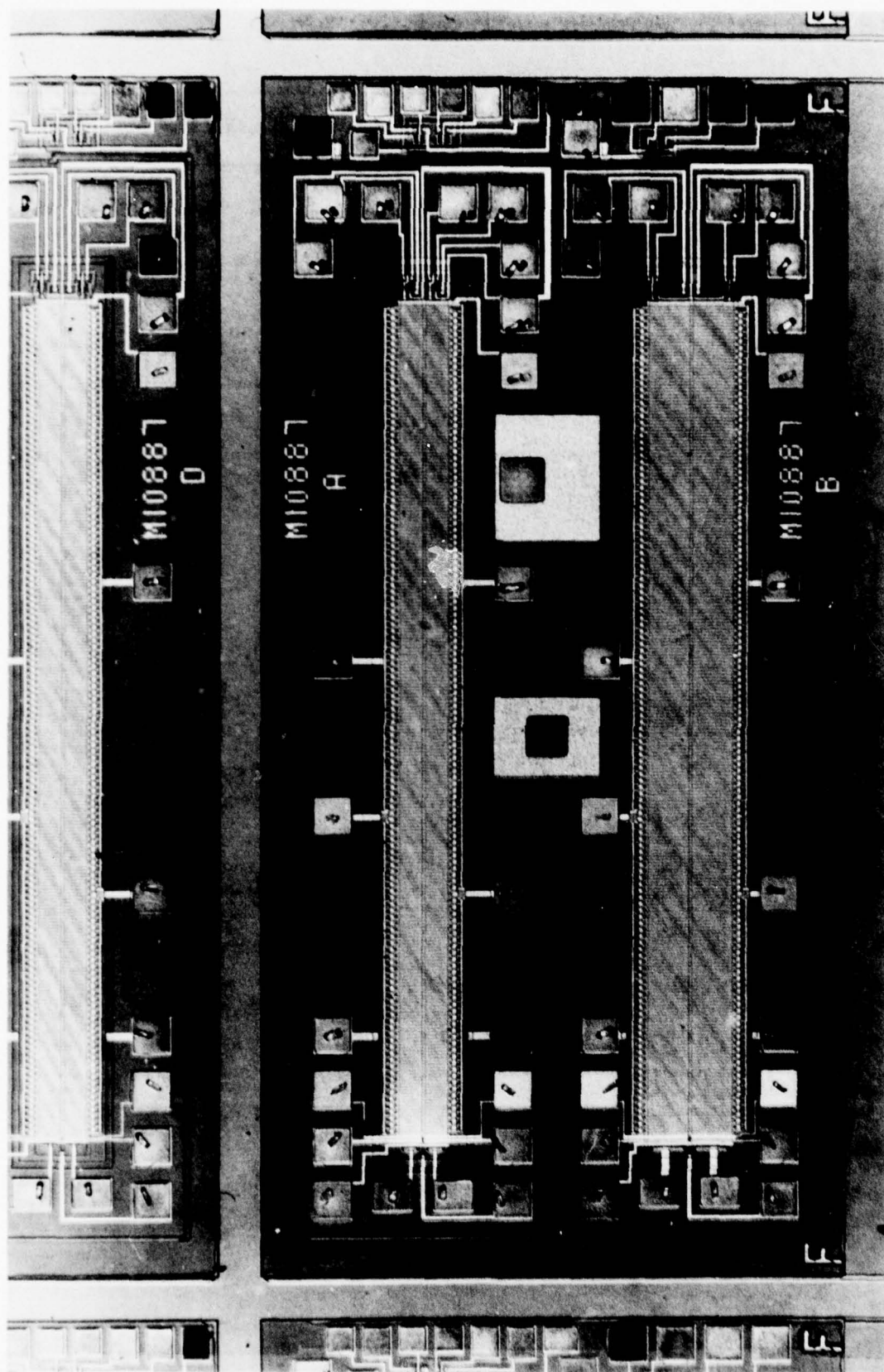Fig.5 Peristaltic CCD — cross section (in direction of charge propagation)
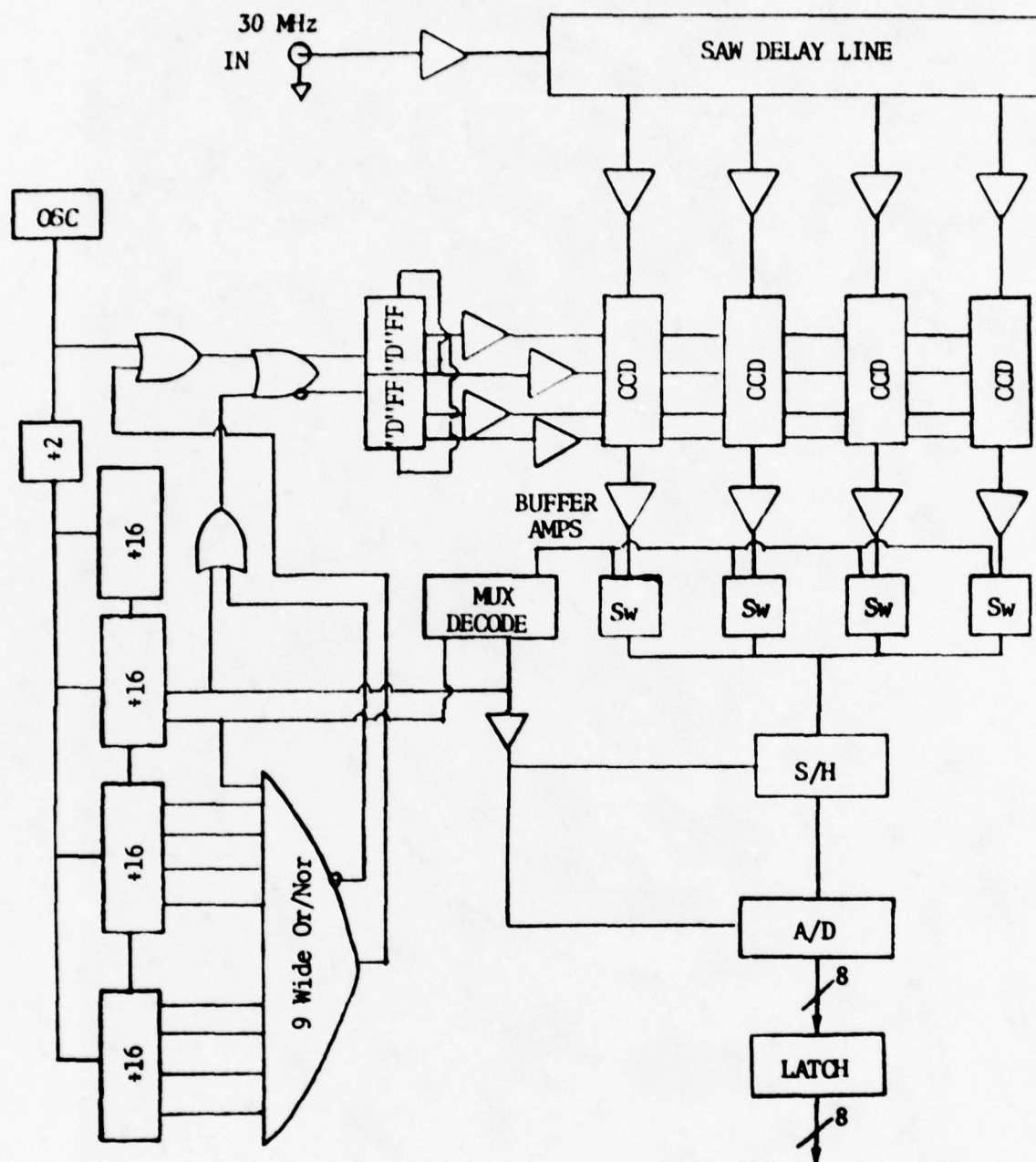
Fig 6 PCCD chip design

Fig. 7   Multiplex SAW/CCD hybrid processor block diagram

A WIDE BANDWIDTH CCD BUFFER MEMORY SYSTEM

K. Siemens, R. W. Wallace, C. R. Robinson
Bell-Northern Research Ltd.
Ottawa, Canada

ABSTRACT

A prototype system has been implemented to demonstrate that CCD's can be applied advantageously to the problem of low power digital storage and particularly to the problem of interfacing widely varying data rates. 8K-bit CCD shift register memories were used to construct a feasibility model 128 K-bit buffer memory system. Serial data that can have rates between 150 kHz and 4.0 MHz can be stored in 4K-bit, randomly-accessible memory blocks. Peak power dissipation during a data transfer is less than 7 W, while idle power is approximately 5.4 W. The system features automatic data input synchronization with the recirculating CCD memory block start address. System expansion to accomodate parallel inputs or a greater number of memory blocks can be performed in a modular fashion. Since the control logic does not increase proportionally to increases in memory capacity, the power requirements per bit of storage can be reduced significantly in a larger system.

1. INTRODUCTION

Data recorder systems in spacecraft have generally been configured to store serial data in tape recorders. Being electromechanical in nature, these systems place limitations on data accessibility, access speed, data rates, reliability, and power. Data may be lost at the start of a transfer since a tape transport mechanism can not respond instantaneously. Solid state memories that are now feasible for airborne and spaceborne applications can surpass the performance and reduce the limitations and deficiencies of electromechanical data recorders. Long-term bulk storage of data in spaceborne applications is expected to utilize magnetic bubble memory systems as a replacement for tape recorders. This type of memory overcomes any objection to moving parts, but it still presents data rate and access speed limitations. It is proposed that these latter problems be overcome by a system using a buffer memory for short-term storage and as an interface between data sources, that operate at a variety of clock rates, and a bubble memory for long-term storage.

A 128 K-bit (1K = 1024) CCD buffer memory system (SIEMENS, K.H. et al .....1976) was developed at Bell-Northern Research (BNR) to demonstrate the feasibility of using charge coupled digital memory devices in a configuration which, used in conjunction with bubble devices, would form a totally solid-state, low-power memory that would have capabilities in excess of the tape recorder systems it would replace.

The shift register configuration of CCD memories is compatible with the requirement to store sequentially arriving data and the shift register storage within bubble memories. The CCD's can accept data, with no loss of bits, over a wide range of clocking rates and transfer this data to the bubble memory at its relatively slow clocking requirement. Since the CCD buffer memory has been designed with a randomly accessible block structure, on-board data processing becomes feasible prior to transfering blocks of data to long-term storage. This feature is difficult to attain with tape recorder systems. The CCD memory block length which is small relative to typical bubble memory shift lengths means that data processing from the buffer storage is easier than from the long-term storage memory. Following extensive testing of the prototype CCD buffer memory by a custom-built memory system exerciser (SIEMENS, K.H. et al ......1976), the unit was delivered to NASA Langley Research Center for further evaluation.

In the following discussion of the buffer memory, it is demonstrated that a CCD shift register memory can be flexible in terms of the data rates it can handle, the lengths of serial data streams that can be stored, and has the capability of access to any set of the memory blocks that make up the structure of the unit. In addition to a description of the system operation, further development possibilities of this type of buffer memory and potential applications of large-scale CCD shift register memories will be discussed.

2. CCD BUFFER MEMORY

The buffer memory system shown in Fig. 1 was designed to evaluate the feasibility of developing a larger scale CCD serial memory that would have capabilities that could be adapted to a variety of applications. The feasibility model incorporates a number of features or design aspects that could be expanded or deleted to suit the particular requirements of a number of potential shift register memory applications. In the following section, the requirements of a buffer memory are explained, the capabilities and features of the feasibility model are outlined, and the system operation is described.

2.1. Design Considerations

A buffer memory system is typically used as an interface between systems that operate at different clock rates. If the buffer were configured in a system as shown in Fig. 2, the buffer and a bulk storage medium such as a bubble memory would form an efficient and versatile tape recorder replacement for airborne and spaceborne applications. The buffer memory could accept serial data streams from a number of data sources and at rates in excess of the capability of the bubble memory. After the data has been captured, it could be transferred to the bulk storage at its own clocking rate.

The combined short-term and long-term memory units offer extended capability compared with a simple tape recorder system. Data need not be sent directly from the buffer to the long-term storage, but

rather to a computer which could perform on-board data processing. The processing could range from transform processing, or data analysis, to a simple decision as to whether or not the data received from sensors is worth storing in the bubble memory. Data could then be routed back from the computer, through the buffer, and on to the bulk storage. The computer could also be used to change the order of blocks of data that are to be transmitted. In these ways, storage space in the bubble memory could be used more efficiently.

Access to a portion of the data held in the long shift registers of a bubble memory is facilitated by dumping back into the buffer memory where it can be processed by the computer. A further use of the arrangement in Fig. 2 is to provide a greater range of data transmission frequencies. Data can be transmitted directly from the bubble memory at only a relatively low frequency. Data held in the buffer can be sent directly to the data transmitter at a high rate. Data could also be transferred at a low rate from the bubble memory to the buffer for re-transmittal at a higher frequency.

With the above configuration in mind, more specific design considerations for a buffer memory must be made before a prototype system can be implemented. Most system requirements are based on the type of data and the manner in which it is to be handled. Thus, the factors affecting the buffer memory input requirements are the data format and data rate. Since the start of a data transfer may not be known before the arrival of the first bit or word, the buffer must be capable of accepting data at any time and with no loss of information.

Output data must have a format that is similar to the input data. Data rate must be variable but independent of the input rate; that is, data written at one rate may later be read out at another rate. Whereas inputs must be captured instantaneously, a certain amount of latency time can usually be tolerated before the data output becomes valid.

2.2. Feasibility Model Capabilities

The feasibility of utilizing charge coupled devices as the storage medium in a buffer memory has been investigated using the Bell-Northern Research 8K-bit CCD (ROSENBAUM, S.D. et al .....1975). These devices consist of 32 dynamic, recirculating shift registers of 256 bits each. The shift register nature of CCD's is compatible with the serial data format of the input data and the bubble memory storage. A single channel with separate input/output is sufficient to evaluate a serial memory. This channel can handle either serial bit streams, serial words, or single bits of parallel words. The system can be expanded to store parallel words by having its memory unit duplicated as many times as the number of bits in a word, without increasing control circuitry.

For serial data storage, CCD's offer advantages over other types of solid-state memory such as RAM's in that a minimum of overhead circuitry and wiring is needed for the data management logic. For RAM's, address information for every word of data must be generated. Data management for shift registers is relatively simple since the storage locations for the serial data sequences are inherently ordered within the register. Address controls for locations within a shift register are *not needed*. Generally, CCD systems can operate at lower power than RAM memories for the same memory capacity and clock rate requirements. This feature is of particular importance in spacecraft applications. A number of design techniques to minimize the power requirements of the CCD buffer memory system will be discussed. At present, the prototype model dissipates 5.4 W in its idle state and 7 W during a data transfer at its maximum rate.

Since the memory is comprised of shift registers and must be capable of buffering serial data bursts, the memory structure is block oriented. A block size of 4096 (4K) bits was chosen to encompass more than the 256 bits of one shift register in order to study any data synchronization or addressing problems, and data rate limitations. 32 data storage blocks were used for a total buffer memory capacity of 128 K-bits.

The 4K-bit block size does not impose a restriction on the length of a data burst that can be stored. The memory has been designed to accept data from one bit to the length of any remaining memory capacity. Each new burst of data, however, will begin storage in the next empty block. Access to the memory blocks is made flexible by a number of options. Any single block may be accessed at random during either a WRITE or READ operation. More than one block can be accessed, since any number of blocks in any combination can be formed into an ordered set. For example, if the controls are set to make the entire memory accessible during a WRITE operation, data will be stored initially in the first empty storage block. If the input data stream exceeds a block length, the next portion of the data will be stored in the next empty block: blocks already containing valid data will be skipped.

Not all of the memory blocks need be made accessible for a data transfer. In a WRITE operation, only the empty blocks within a selected set will be filled in order. Conversely, during a READ operation, only the blocks which contain valid data will be read in order, out of the set that has been selected. Two methods of read-out are permitted, non-destructive or destructive. With the first option, the data in the memory block is retained. With the latter option, the memory block is made available for new data storage immediately after the block has been read. Data is always stored as a first-in, first-out serial data stream within any block. Since the buffer memory is used primarily to interface systems that operate at different clock rates, the WRITE and READ operations can not be performed simultaneously.

The buffer memory can transfer over a broad range of data rates, between 150 kHz and 4.0 MHz. The rate is dependent on the external clock used to drive the memory. The dynamic nature of the CCD's

requires that data be clocked to keep it refreshed even when the memory is not being accessed. The idle clocking rate is kept between 10 to 20 kHz, regardless of the rate of the clock used to drive the system. With constant data recirculation during the idle state, the start of a memory block may not be immediately accessible when a data transfer is requested. However, when a data input begins, there can be no latency period before data can be accepted. A pre-buffer adapts itself to synchronize input data with the recirculating memory block. During a READ operation, there may be a waiting period up to 1023 clock cycles until the start of the memory block is accessible for a data output. If a readout is requested during a WRITE operation, the data output will become available immediately following the input transfer.

2.3.   Prototype Memory Operation.

In previous sections, some of the design considerations and capabilities of a CCD buffer memory system have been discussed. Here, a general description is given of the internal operation and features of the prototype CCD memory system. The block diagram in Fig. 3 is used as a reference.

In essence, the memory system takes a serial input data stream, routes the data to the first available memory block, and stores the data in successive memory locations. Then upon command, the data in any memory block can be read and presented to the memory system data output terminal. Complications to this basic concept arise from the dynamic nature of the storage medium. An 8-bit counter, included as part of the clock controls, provides a virtual address (VA) for each of the 256 storage locations on a CCD shift register. The VA is incremented with each CCD transfer clock cycle, which is maintained at an idle rate while the system is not being accessed. Data storage must always begin at the start address of any memory block. This start address is determined arbitrarily by the zero state of the virtual address counter, i.e. VA = 0. Since input data may arrive at any time, it is unlikely that the start of the incoming data stream will coincide with the memory block start address. Therefore, the pre-buffer that acts as a variable delay temporarily stores incoming data until it can be written into the CCD's after their start address has been reached.

Further complications to the design are caused by the need to handle data rates in excess of the capability of individual CCD's. Since the 8K CCD is specified to operate at rates from 10 kHz to 1 MHz, 4-way data multiplexing was used to achieve data transfer rates up to 4 MHz. The 150 kHz lower limit for the system was set as an arbitrary design goal to match the expected clocking rates of bubble memories. At the lowest data rate, 4-way multiplexing means that the transfer rate to individual CCD's will be only 37.5 kHz.

A serial-to-parallel converter in the pre-buffer is used to route input data along 4 parallel paths to the CCD's. The CCD's, shown on the memory card in Fig. 4, are arranged in a 4 by 4 array to accomodate the parallel data inputs. Each bank of 4 memory devices is used to store 8 data blocks. When the memory is being accessed, 4 bits of data - 1 for each device in a bank - are transferred simultaneously with every CCD clocking cycle. Similarly, during a readout operation, the memory waits until the start address of the first memory block to be read is reached, then 4 bits at a time are transferred to a parallel-to-serial converter in the data output buffer. In this way, the output is presented as a first-in first-out serial data stream at the master clocking frequency.

The block selector logic plays a key role in the data routing process. In a WRITE operation, only empty blocks may be accessed; in a READ mode, only blocks containing data may be accessed. Versatility of memory block access is provided by an iterative array of block selector cells and block flags. During a WRITE operation, a block selector cell examines whether the block has been selected for a possible transfer, and the status of its own and lower order block flags. If no lower order blocks are available, it asserts its "block selected" signal which is decoded to form the CCD shift register addresses and chip enable signals. When data transfer to a selected block begins, its block flag is set, and the block selector searches for the next higher-order block to be accessed, so that the next storage block is available immediately. During a readout operation, the block selector indicates only memory blocks that contain valid data. An option has been provided to retain data blocks that have been read or to reset the block flags thereby freeing the blocks for fresh data storage.

Sequence and clock controls are designed for the CCD buffer memory to operate in 3 modes; idle, transfer set-up, and data transfer mode. An example of the timing of these modes is illustrated in Fig. 5 for an input transfer of 2 1/2 data blocks. When the system is in idle, the CCD's are clocked at about 10 kHz to keep data refreshed. A low idle rate helps to minimize power consumption. The system uses gated external clock pulses to ensure that CCD recirculation even during the idle state is synchronized with the master clock.

When a request for data transfer is initiated at time A, the system goes into a transfer set-up mode. For a WRITE operation, the presence of valid data is detected by a Data Input Available signal which is asserted for the duration of the input burst. All memory blocks then recirculate at the data rate until the ZERO virtual address is reached at point B. Data is then transferred from the pre-buffer to the first available memory block. In order to reduce system power during a transfer, clocking of all non-accessed banks of CCD's is reduced to 1/4 the CCD transfer rate. At this rate, all non-accessed shift registers will recirculate an integral number of times during a block transfer so that the virtual address of all registers will again be synchronized.

When the input data stream ends at C, the tail end of the burst still resides in the pre-buffer, so that the CCD transfer continues for the equivalent of the initial transfer set-up time, up to time D. If a memory block is not yet full at the end of that time, the system continues in the transfer mode to allow the rest of the block to fill up with ZERO's. The system will then revert to its idle state at E.

The operating sequence for a readout operation is similar to the WRITE sequence. After the blocks to be read out have been selected, the READ command is asserted. The system goes into the transfer set-up mode until the start address of the first block has been reached. Then, successive, complete blocks of data will be read. The system provides a Data Output Available signal for the duration of valid output data.

The buffer memory is a low power system that dissipates 5.4 W in the idle state. Since CMOS logic and MOS devices that have high capacitance are used, the power consumption varies with the input clock frequency. The peak power of approximately 7 W occurs during the WRITE operation at the maximum data transfer rate. Measured power dissipation curves for WRITE and READ operations for the full operating range of data rates are shown in Fig. 6. For the complete system, the average power dissipation per bit of storage is less than 50 μW. In the following section, it will be shown that the average power per bit can be reduced significantly if the prototype buffer memory is modified for other applications.

3. ALTERNATE CCD BUFFER APPLICATIONS

The present memory system is a prototype feasibility model of limited capacity and contains features that can be simplified or augmented for any particular application. As other CCD's become available with a variety of shift register lengths, memory configurations, and data bandwidths, restrictions on the buffer will change. These devices will permit a greater degree of freedom in selecting levels of multiplexing, data rates, or block sizes.

A considerable amount of logic was included in the prototype to provide manual control to facilitate the system evaluation. The logic circuitry, contained mainly in the block selector, could be replaced by computer or microprocessor control to reduce the peripheral circuitry. Although, for an expanded version of the buffer, microprocessor control should be considered, it would have consumed too much power to meet the design goals in the prototype model.

The microprocessor could store the status of the data blocks and, if necessary, some pertinent information relating to the material content of each data block. Recall can be affected, under program control, according to the type of information required, the data source, or even the time of data capture. A substantial portion of the input pre-buffer could be eliminated if the microprocessor were to store the CCD virtual address concurrent with the start of the block transfer. There would be no need to wait for the CCD ZERO virtual address before starting the CCD input transfer. Output transfers would begin when the block virtual address matched the stored start address.

The data storage size does not have to remain constant if the system is under microprocessor control. The buffer could adapt the block length to the input data burst for more efficient use of storage space.

It may be noted that with either block selector logic or computer control, the peripheral hardware does not increase proportionally to increased memory capacity. This is particularly true if duplicate memory units are added for parallel word storage. Therefore, in a large memory system the average power per bit could be reduced substantially until it approaches the limits of the CCD and driver dissipation. This is expected to be at most a few microwatts per bit.

4. FURTHER APPLICATIONS OF SERIAL MEMORY SYSTEMS

The 128 K-bit CCD buffer memory system has demonstrated the feasibility of using CCD's to capture sequentially arriving data streams at varying data rates and to store the data as blocks of information for subsequent transmission. The prototype system could be used as the basis of the design of a substantially larger, more flexible buffer memory of similar configuration and under microprocessor control. The same model could be reduced to a simple tape recorder replacement with little versatility of access and with no pre-buffer to prevent a minor loss of data at the start of transmission.

There are a number of other areas in which CCD memories should be considered as an option to other forms of solid-state memory, such as RAM's. Serial data, by its nature, does not always need the constant address supervision that must be used for RAM's. Generally, the only information needed is the start of the data stream, whether the data is partitioned into blocks or frames, and the length of a block or data burst. Typical applications for which serial memories could be used are disc replacement memories (SIEMENS, K.H. et al .... 1975) and video frame memories. The video frame memory is probably the application that can be exploited most easily since the basic requirement is a very long shift register with no random block access requirement. CCD's are now available (ROSENBAUM, S.D. et al .... 1976) which will operate at video rates. Another potential application that shows good promise is in the area of synthetic aperture radar (SAR) signal processing. A key component of a SAR system is a corner-turning memory array. A technique has been found (SIEMENS, K. et al .... 1977) using CCD's that could be developed into a very efficient corner-turning memory that is large enough to provide high resolution and to permit the SAR system to operate in real time for on-board applications.

The buffer memory system described in this report is one of the first system applications for CCD memories. Other applications that are now considered feasible have been mentioned. It is expected that as larger capacity CCD's with a variety of configurations become more generally available, a number of new uses of large-scale, solid-state serial memories will be developed.

REFERENCES:

ROSENBAUM, S.D., CAVES, J.T., 1975, "8192-Bit Block Addressable CCD Memory", IEEE Journal of Solid State Circuits, Vol. SC-10, October 1975, pp. 273-280.

ROSENBAUM, S.D., CHAN, C.H., CAVES, J.T., POON, S., WALLACE, R.W., 1976, "A 16 384-Bit High Density CCD Memory", IEEE Journal of Solid State Circuits, Vol. SC-11, February 1976, pp. 33-40.

SIEMENS, K., CHOWANIEC, A., WALLACE, R.W., 1977, "A Feasibility Study on Utilization of CCD's for a SAR Signal Processing System", Report prepared by Bell-Northern Research Ltd., March 1977.

SIEMENS, K., ROBINSON, C.R., MASTRONARDI, J., 1975, "A Fast Access Bulk Memory System using CCD's/A Recorder Buffer Memory using CCD's", Proc. 1975 Int. Conf. on the Appl. of Charge-Coupled Devices, San Diego, October 1975, pp. 429-434.

SIEMENS, K.H., WALLACE, R.W., ROBINSON, C.R., 1976, "A 128 K-bit CCD Buffer Memory System", NASA CR-145020, Report prepared by Bell-Northern Research Ltd., July 1976.

Fig. 1   128 K-bit CCD Buffer Memory System

Note: Arrows Indicate Data Routing.
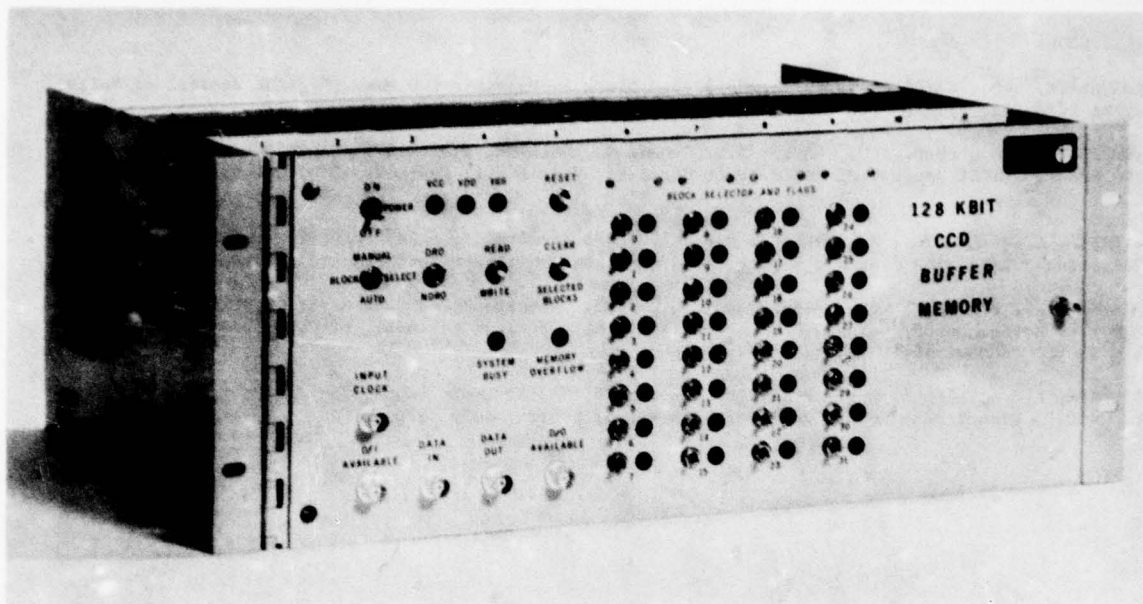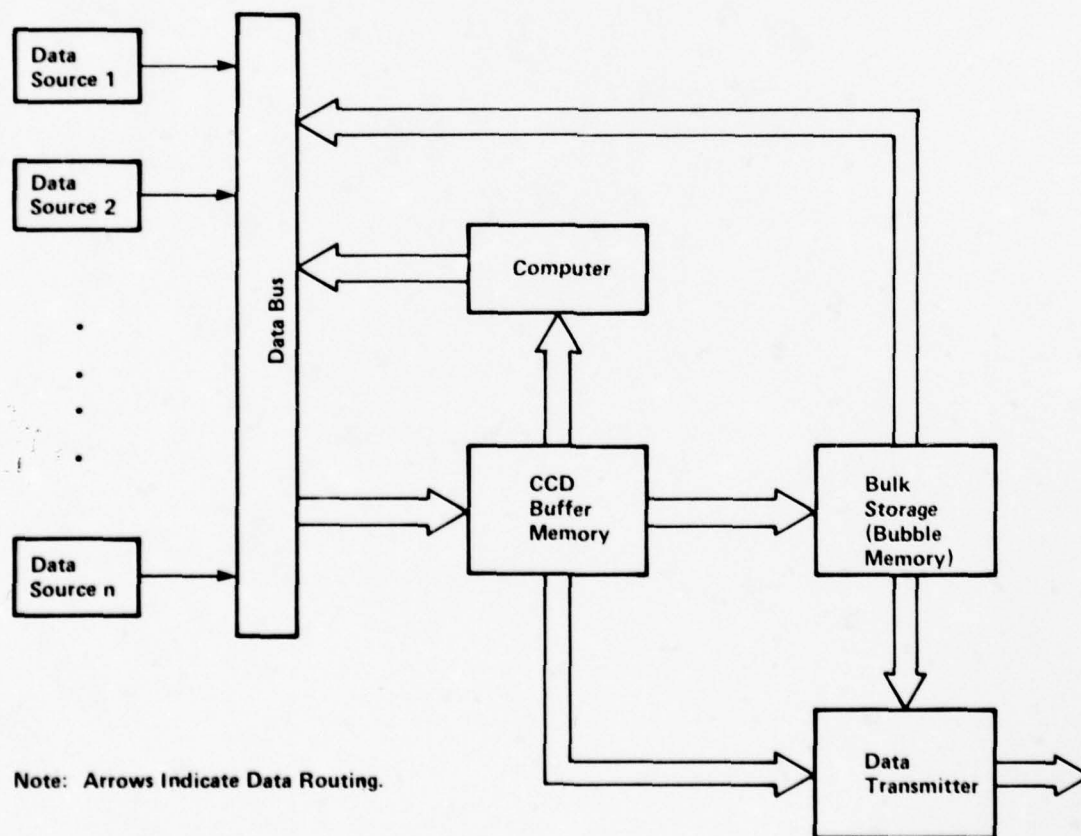
Fig. 2   Information Flow in Buffer Memory System
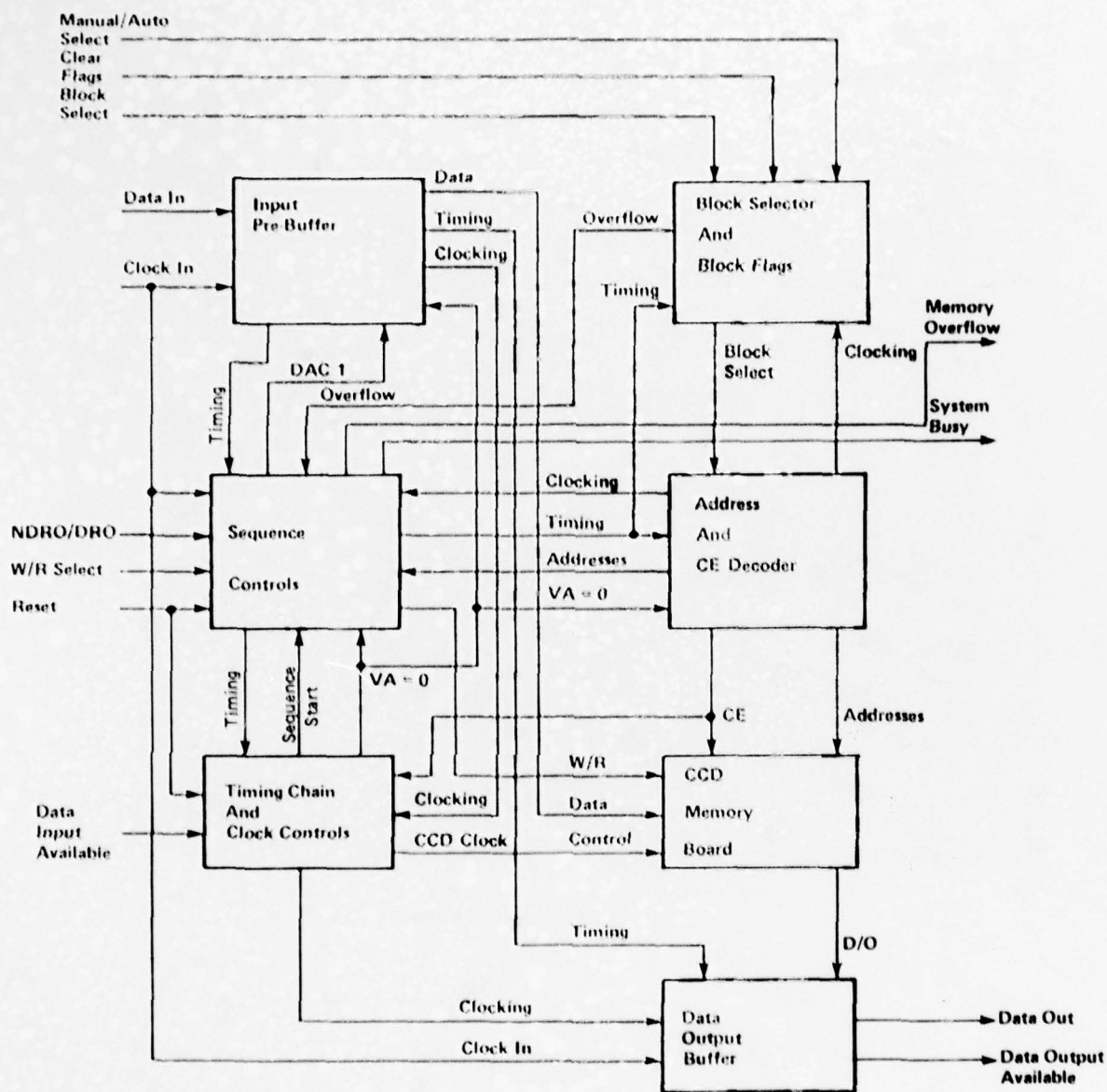
Fig. 3   Block Diagram of Buffer Memory System
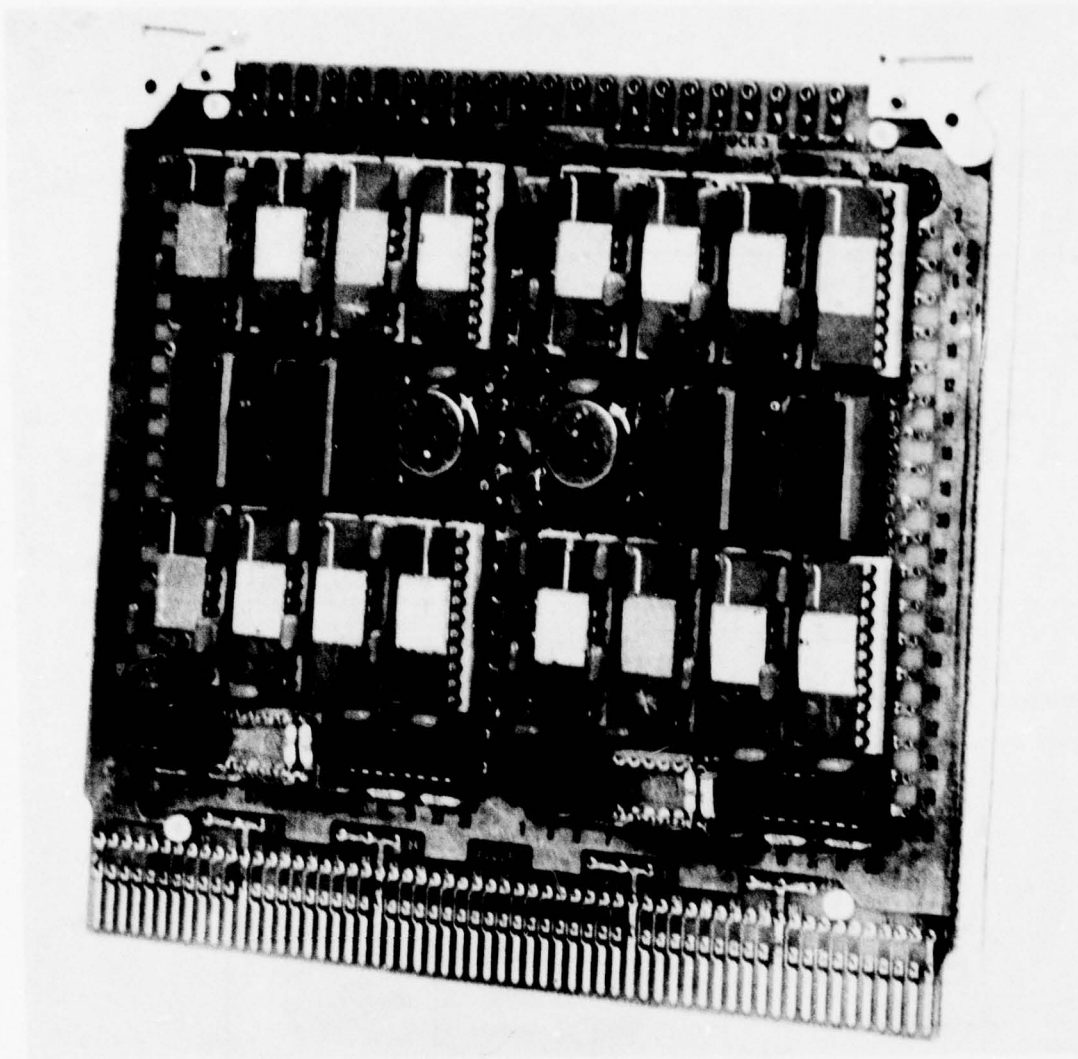
Fig. 4 CCD Memory Card
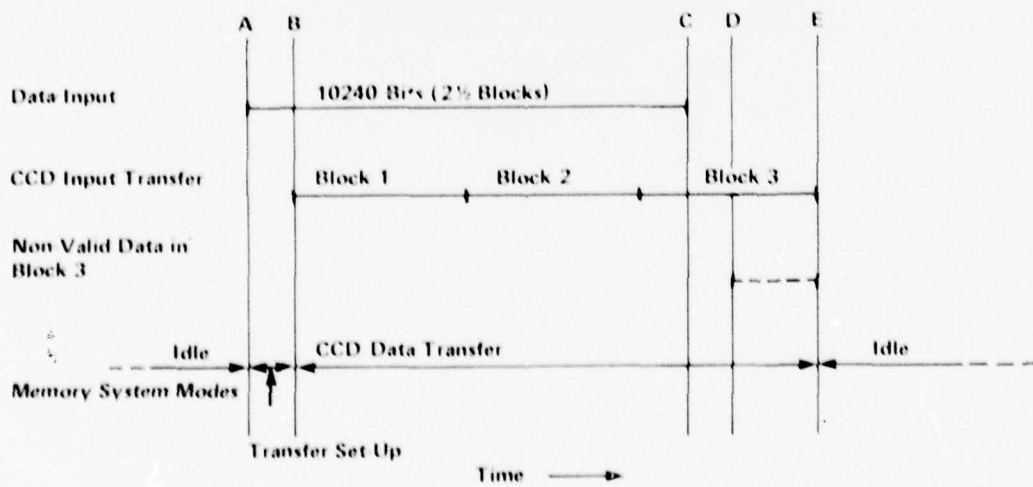
Fig. 5 Example of Operating Sequence during Input Transfer
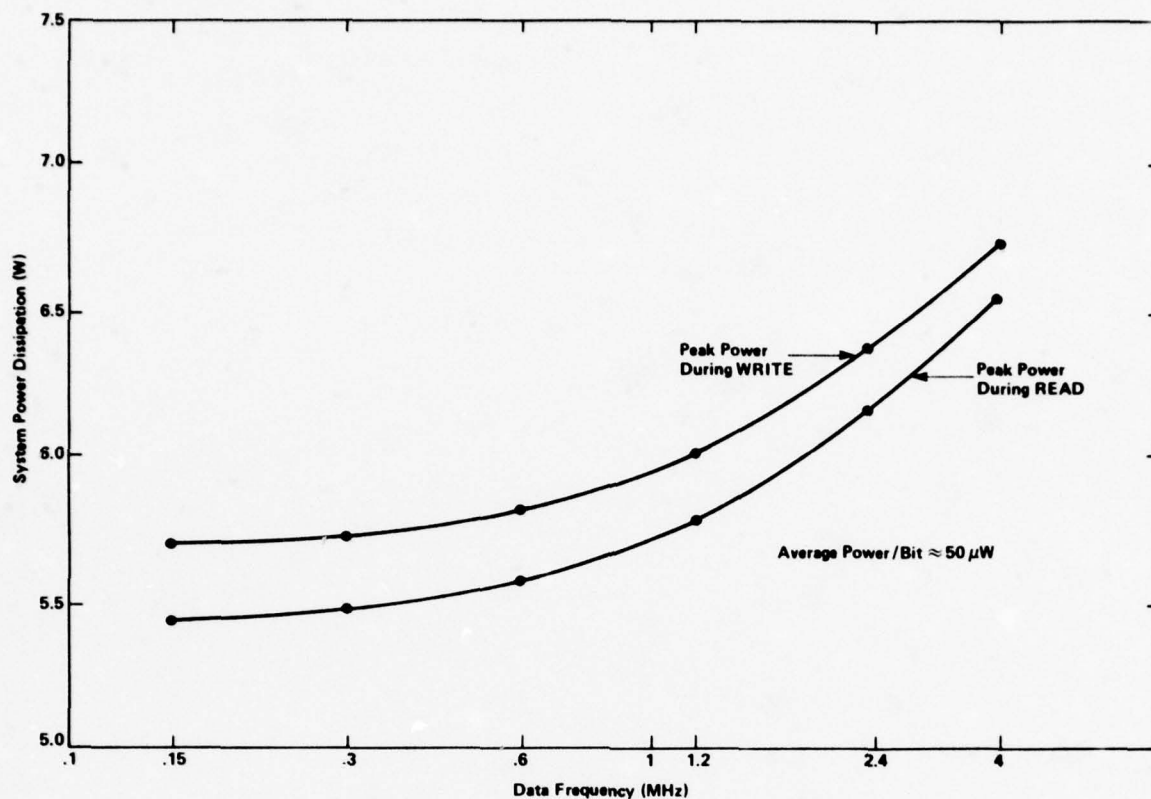
Fig. 6   System Power Dissipation

DISCUSSION

J.J.Stapleton

1 — Since you frequently repeated that your 128-K system was a feasibility model, I wonder what problems you experienced? Do you think digital CCD's will be soon in production? Finally, were all CCD's on the same power supply?

Author's Reply

1 — a) The feasibility model was designed to ascertain the operating limits of both the CCD's and the buffer system configuration. Emphasis was placed on techniques to minimize power consumption. Since this meant operating the CCD's near their lower voltage limit, the operating margins were reduced. Clock capacitance of about 400 pF meant that one clock driven could operate only two devices. BNR's new, larger-capacity CCD's, however, have a clock capacitance of only 70 pF.

1 — b) We are not acquainted with the future marketing strategies for digital CCD's by semiconductor manufacturers.

2 — Yes, all CCD's were on the same power supply, that is, $V_{DD} = + 10\ V_{DC}$, $V_{BB} = -3\ V_{DC}$.

A MICROPROCESSOR CONTROLLED
ELECTRICALLY PROGRAMMABLE TRANSVERSAL FILTER

A. CHOWANIEC, R. WALLACE
Bell-Northern Research, Ottawa

M. A. COPELAND, Y. A. HAQUE
Department of Electronics, Carleton University, Ottawa

ABSTRACT

This paper describes a real time electrically programmable transversal filter which promises to have considerable potential for sampled analog signal processing and presents results obtained using the filter in combination with a microprocessor. The tap weightings are controlled by a digital signal and a microprocessor is used to exercise the filter by allowing rapid changes to be made to the filter characteristic.

The filter is based on a novel functional multiplying structure in the standard silicon gate MOS process which has previously been described. The analog multiplication of signals is achieved by digitally selecting the application of controlled analog signal voltage on the semiconductor inversion layer side of an MOS capacitor – a process that is linear and free of signal interaction (crosstalk) between taps. The sampled analog signal is stored at fixed sample and hold storage sites on chip, and digitized analog tap data is loaded externally and moves relative to the chip. The digitized tap data in two's complement format controls the selection of seven binary area-ratioed MOS capacitors per tap position for seven-bit tap accuracy. The multiplying capacitances are commoned across the chip to achieve summation without the necessity for amplifiers.

A microprocessor is used to manipulate and load the tap data. It allows complete sets of tap weights to be loaded and then recirculated giving the filter a stable frequency response. The processor can be used to store tap data for numerous filter transfer functions which can then be loaded. Alternately, the keyboard can also be used to access and modify individual taps in an existing data pattern.

A programmable filter of this type can find wide application in advanced signal processing system and some discussion will be presented on its applicability to areas such as adaptive filtering for communications and matched filtering for spread spectrum modems.

1. INTRODUCTION

With the growth of charge transfer device (CTD) technology the use of transversal filters with fixed weighting coefficients has been established in a variety of signal processing applications. There is increasing interest in programmable structures of this type where the weightings can be altered in real time. This paper considers one possible implementation of a monolithic programmable structure which does not employ CTD technology but which promises good device performance (HAQUE, Y.A., and COPELAND, M.A. .... 1976)

In this structure the analog input signal is sampled and stored on fixed sample and hold storage sites on chip. The tap weightings are introduced on chip in digital form and control the selection of area-ratioed MOS capacitors associated with each sample and hold stage or tap position. These capacitors form the basic multiplying structure of the filter with signal in terms of voltage rather than as charge in CTD designs. This has the advantage of eliminating cross-talk between taps. The use of a low-impedance voltage source minimizes the effect of various non-linear capacitances (HAQUE, Y.A., and COPELAND, M.A. .... 1977).

The filter is implemented with a single level MOS/LSI process and is interfaced to a microprocessor which feeds the digital tap data to the chip and allows real-time manipulation of the filter characteristic.

2. CHIP ORGANIZATION

A block diagram of the on-chip circuitry is shown in Fig. (1). The tap weightings are specified using 7-bit data words employing two's complement coding. The words are clocked through shift registers to each tap location where they control the switching of the analog signal samples on to the sensing surface of the seven ratioed capacitors. The ring counter provides a pulse which gates the analog input signal into the sample and hold at that tap position only once every N clock cycles for an N tap device.

The details of each tap position have previously been described (HAQUE, Y.A. and COPELAND, M.A., ... 1976) and are briefly summarized here with the aid of Fig. (2) which shows one of the seven capacitors at one tap location. The sensing electrode is maintained at a sufficiently high voltage to form an inversion layer which is then addressed by the analog signal voltage (V_S) and the reference voltage (ground) via the switches $\emptyset_2$ and $\overline{\emptyset2}$. The $\emptyset_2$ switch is controlled by one bit of the digital word.

The change in voltage at the semiconductor surface as it is switched from reference voltage to signal voltage or vice-versa is sensed by the electrodes. The operating waveforms are shown in Fig. 3 for the present structure which consists of two multiplexed channels with the same filter function.

Consider the channel operating with the $\emptyset_2$ clock. When $\emptyset_g$ is on the sensing electrode it is constrained to V_{dd} and with $\emptyset_2$ on the sensing surface is at the signal voltage. Clock $\emptyset_g$ next goes off followed by $\emptyset_2$ at which time the sensing surface is reset to ground and a negative going transition is sensed at the output.

3. SYSTEM ORGANIZATION

The peripheral circuitry required to operate the filter chip is shown in Fig. 4. The amplifier at the input is used to eliminate the non-linearity and gain variation of the on-chip buffer amplifiers. Data from the microprocessor is routed to the chip under control of a 'load' signal. If a load pulse is present the data is fed to the chip. If it is absent data is routed from the output of the chip back to the input. In this mode a given set of tap data can be continually recirculated to give a fixed filter characteristic.

Figure 5 shows the microprocessor interface which controls the loading of the tap data into the filter. The interface does not operate in a handshaking mode but in a switched control mode so that once loaded the system is independent of the microprocessor. The interface appears to the microprocessor as a RAM. The address decoder produces a valid output when it recognizes data meant only for the interface. The load counter generates a pulse once for every N clock cycles where N is the number of stages in the filter (32 in this case). This pulse is the ring counter for the chip and also acts for the interface as a virtual zero address (VZA) to indicate the first tap position. This signal gates the clock which is used to drive the tap address counter. This counter then counts down from the low address input from the microprocessor towards zero at which time it provides the filter load signal at the correct time slot for that filter tap. As far as the interface is concerned the filter acts as a shift register memory. The mode control ensures that the gate signal is only provided when a decoded write occurs and is turned off once a load signal is generated.

The interface requires up to N-1 clock cycles to load data for one tap. As a result of this a software timing loop has to be used to ensure that the microprocessor does not present data faster than the interface can accept.

The interface allows the user, via the microprocessor input keyboard, to load a complete set of taps and/or a new filter characteristic to change individual taps and modify the response.

4. EXPERIMENTAL RESULTS

The magnitude response of a filter loaded with coefficients for a linear phase low pass design is shown in Fig. 6. Four cascaded chips were used to give a total of 32 taps. The clock frequency was 50 KHz and the first aliased passband is shown at the right of the figure. Fig. 7 shows a similar magnitude response but in this case the coefficients are for a minimum phase design.

Figures 8 and 9 show the effect of convolving the input signal with a rectangular tap weight distribution. In figure 8 the first sixteen coefficients are set to unity, the rest to zero. In Figure 9 the first sixteen are unity whilst the remainder are unity and negative.

5. DISTORTION AND FIXED PATTERN NOISE

The primary limitation of this device is the reduction of its usable dynamic range due to fixed pattern noise (FPN). The output of the filter with only a dc bias as its input consists of spectra at 1/N (normalized frequency) and its multiples for an N tap filter and is defined to be the fixed pattern noise (shown in Figs. 6-9).

An N tap filter can be considered to be a structure with channels multiplexed in the time domain (HAQUE, Y.A., and COPELAND, M.A., ... 1977). Any deviation in channel characterisitics will show up as fixed pattern noise (ROY, D., et al. .... 1977). The specific sources of channel characteristic deviation are: gain variation of the channels, variation of capacitive feedthrough from multiplexing clocks in the different channels and tap weight (which spends equal time in all channels) variation across the channels. This effect also causes interaction products between the input signal frequency and the multiplexing clocks. FPN can be considered to be a special case of this interaction product: i.e., interaction with dc.

Fixed pattern noise in the present chip design has been greatly suppressed by the use of feedback linearization, which corrects the gain variation across channels. The effect of capacitive feedthrough in the channels has also been suppressed using an optimum signal input and sensing scheme (HAQUE, Y.A., and COPELAND, MA., ... 1977). The effect of tap weight errors could not be corrected and is believed to be a major remaining cause of FPN.

The tap weight errors cause FPN as follows.
The tap weight at every location consists of binary area ratioed capacitors which are suitably interconnected under control of digital switches, fed with the binary coding for the tap weight. These binary codings move relative to the tap locations in a periodic manner. Due to variation of the size of the area ratioed capacitors at different locations (channels), a time varying error with period NT, (T being the bit period of the filter with N taps) can be associated with every tap weight.

At present, 45 dB dynamic range (with less than 1% harmonic distortion) has been obtained for typical filters on 3 ohm cm substrate material (the dynamic range is defined to be the difference between the passband maxima of the filter and the maxima of any FPN spectra with less than 1% harmonic distortion in the filter passband maxima). The programmed filter function strongly affects the

dynamic range, since the FPN due to gain variations and capacitive feedthrough in the channels is filtered by the filter transfer function. In addition, the nature of the tap weight error would depend on the programmed function. Fig. 10 shows a 60 dB dynamic range for a rectangular window (tap weights are all one's in the time domain for a 96 tap filter) which eliminates the effect of tap weight errors. The effect of the tap weight errors causing FPN in this case is eliminated, since the total errors for all tap weights is constant for all time at all locations in the filter (since connections to the area ratioed capacitors are constant). Further improvement in the dynamic range is expected to be obtained by improvement in the chip design (which would suppress the effect of capacitive feedthrough and ground loops) (HAQUE, Y.A. and COPELAND, M.A., ... 1977).

Transfer functions measured for different filters have been found to be surprisingly close in agreement with calculated responses in spite of considerable tolerances in tap weight coefficients. Statistical studies on tap weight errors on the transfer function of fixed tap filters (HAQUE, Y.A., ... 1977) (PUCKETTE, C.M., et al .... 1977) predict less stop band attenuation than we have achieved. The explanation for this effect is an averaging of tap weight errors taking place in the filter during the measurement process. Such averaging reduces the error variance (over that for fixed tap filter) and yields a better stop band attenuation. This effect may also be viewed from the frequency domain: The tap weight may be thought to be comprised of an expected value (which is constant as the tap weight traverses the different tap locations) and a time varying error which is periodic with frequency $1/NT$. An input signal of frequency fs results in an output at frequency fs (due to the constant value of the tap weight vector). The error part of the tap weight can be expanded as a Fourier series with random coefficients. This periodic waveform amplitude modulates the input frequency fs and thus creates sidebands at $fs \pm 1/NT$ (assuming that only the fundamental in the Fourier series expansion is dominant). Thus in order to measure the effect of the time varying tap weight errors on the transfer function, the measurement bandwidth must be at least $2/NT$. Experimentally, this condition was not met because of the small bandwidth of the spectrum analyzer (which creates a long time response, effectively averaging the time varying response in the time domain (2)). The effect of tap weight tolerances is therefore suppressed in the measurement procedure. In general, the time varying nature of the tap weights would yield better stopband rejection (compared to a comparable fixed tap filter) in systems with bandwidth less than $2n/NT$, where n is the number of harmonics (of the fundamental at $1/NT$) carrying most of the energy in the periodic time varying waveform of the tap weight error.

6. SUMMARY

The programmable transversal filter or correlator described in this paper promises to have considerable potential for signal processing systems, particularly as significantly improved performance over the present prototype device can be expected. In a redesigned structure fixed pattern noise is expected to be below 65 dB. Although the present chips are cascadable, and this feature could be retained, it is feasible to implement at least 32 taps on one chip and at the same time add additional bits to the coefficient data word. This is the result of the fact that the basic structure of the device is inherently simple to implement in MOS technology. The size of the MSB capacitor is the *major limiting factor. Although the device clock rates* used in the work presented in the paper are relatively low, the device will operate with clock frequencies in the low megahertz.

One application of such a structure is as a matched filter in spread spectrum modems. The purpose of the filter is to synchronize the system to the incoming code and consequently the filter is dependant on the code. A programmable structure which could be matched to different codes is therefore of advantage. The major limitation is that the code lengths can require devices with many hundreds of taps and therefore large numbers of cascaded devices. In the case of matching to a p-n sequence rather than an analog signal the system could be configured in such a way that it would be sufficient to control the weightings with only one bit of the data word. A device built for this situation would therefore allow a large number of taps to be implemented on one chip. Cascading of devices is again possible although this will exacerbate the FPN problems.

Although a microprocessor has been used here as one way of exercising the filter it is by no means an integral part of the device. Several such filters could be controlled remotely by a processor or alternatively one filter could be combined with a ROM which would store a limited number of weighting coefficients.

Real-time adaptive filtering is also possible if the filter is provided with additional circuitry to carry out the updating of the tap data. The algorithms for adaptive structures have been extensively discussed in the literature and can be implemented with circuitry of varying complexity. In all cases access to the analog signals at the sample and hold stages is required and this in turn implies some form of A/D conversion.

REFERENCES:

HAQUE, Y.A., COPELAND, M.A., International Electron Devices Meeting, Washington D.C., 1976 pp. 27-30.

HAQUE, Y.A., COPELAND, M.A., IEEE Journal of Solid State Circuits, to be published Dec. 1977.

ROY, D. et al, IEEE Trans. Electron Devices, Vol. ED-24, No. 6, June 1977.

HAQUE, Y.A., Ph. D. Thesis, Carleton University, Ottawa, Canada, 1977.

PUCKETTE, C.M., BUTLER, W.J., SMITH, D.A., IEEE Trans. Comm. Technol., Vol. COM-22, July 1974, pp. 926-934.

Fig. 1 Chip organization.

Fig. 2 Detail of one binary coefficient at one tap location.
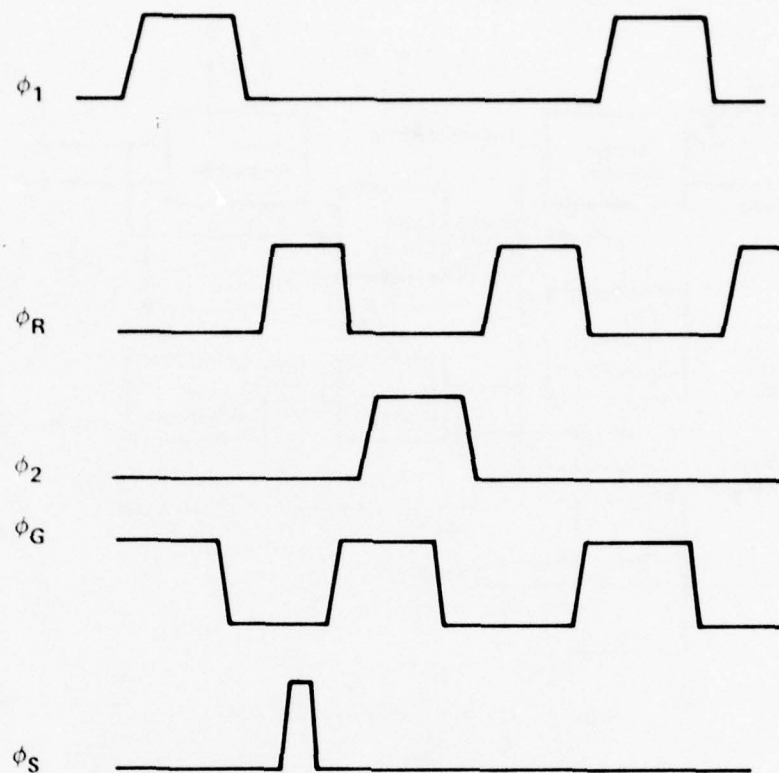
Fig. 3   Operating waveform.
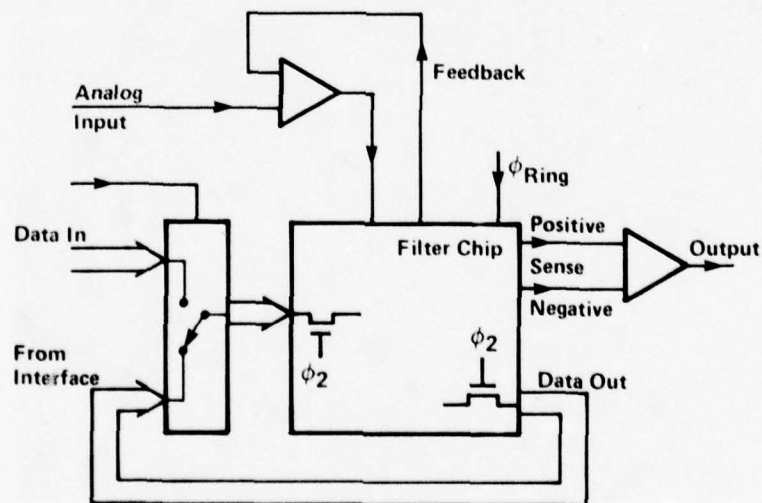
Fig. 4   Peripheral circuitry required to operate the chip.
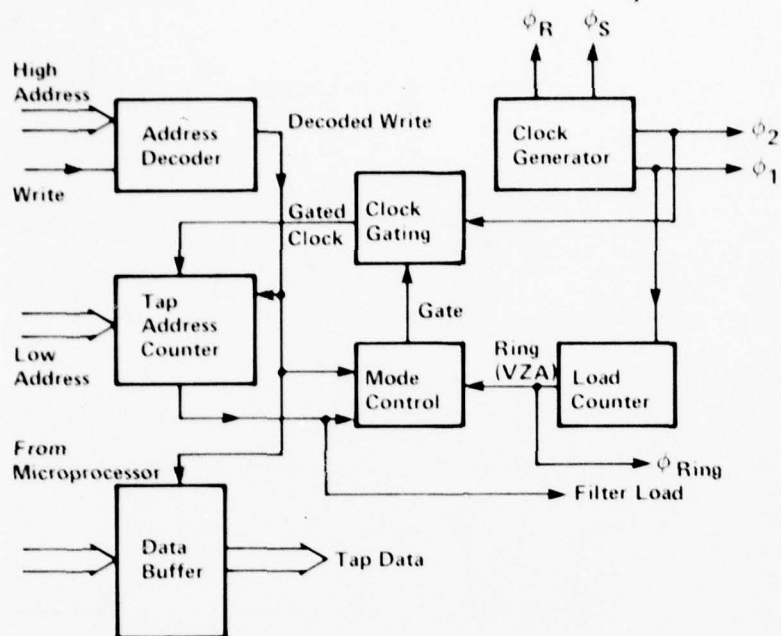
Fig. 5   The microprocessor interface.
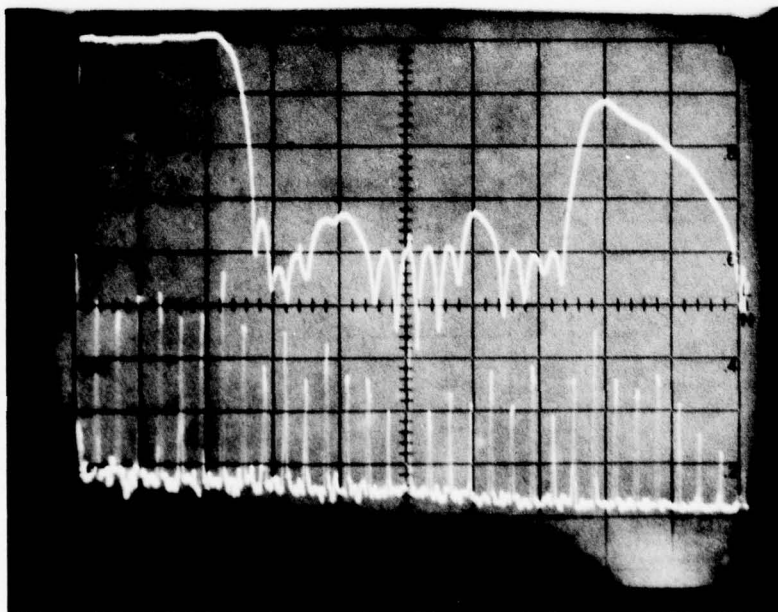
Fig. 6   Magnitude response for low-pass linear phase 32 tap filter design and corresponding fixed pattern noise (lower curve). The frequency is at 50 KHz. The vertical scale is 10 dB/division.
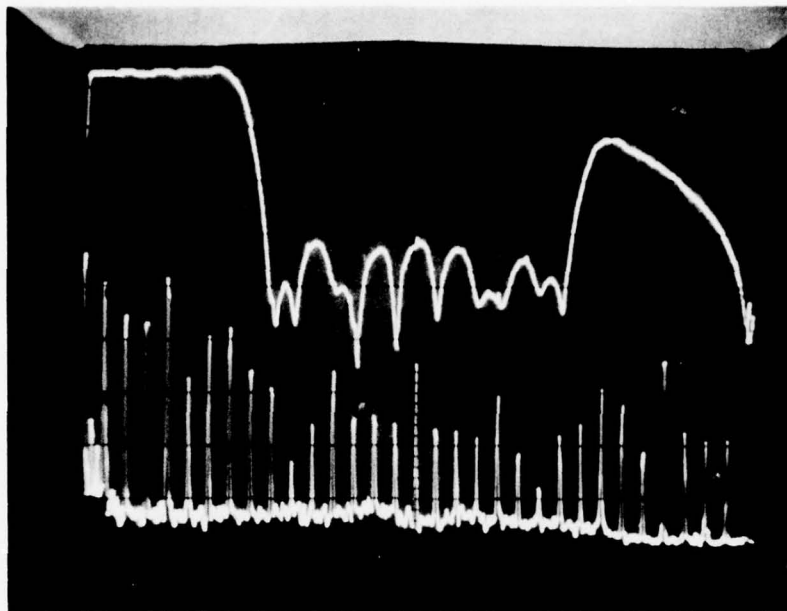
Fig. 7   As above except that coefficients for a minimum phase design are used.
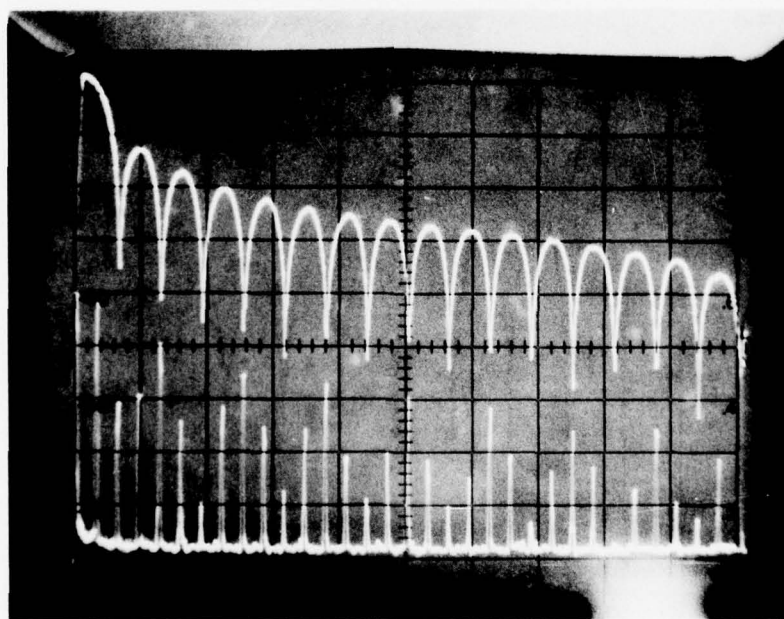
Fig. 8   Magnitude response and fixed pattern noise (lower case) for the filter operating with sixteen unity weighted coefficients followed by sixteen zero weighted coefficients. The clock frequency is 50 KHz.
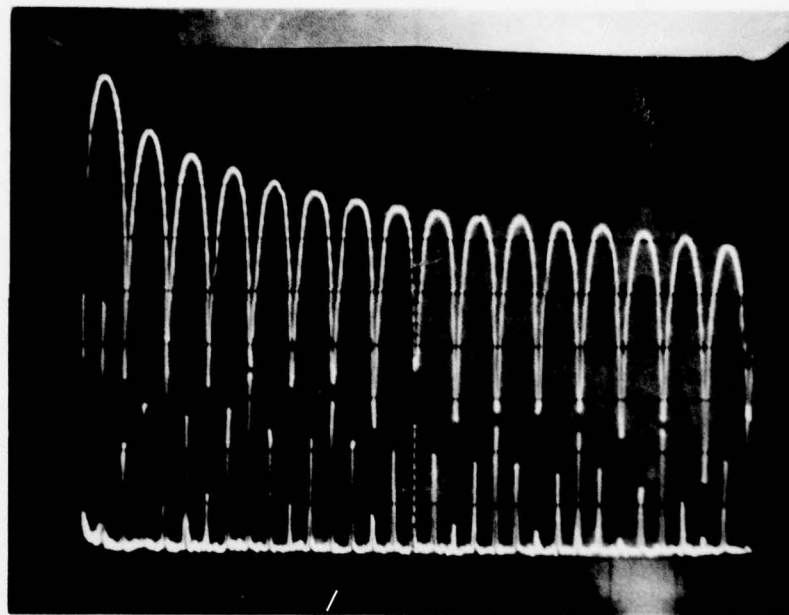
Fig. 9 As above except the zero weighted coefficients are replaced by unity, negative coefficients.
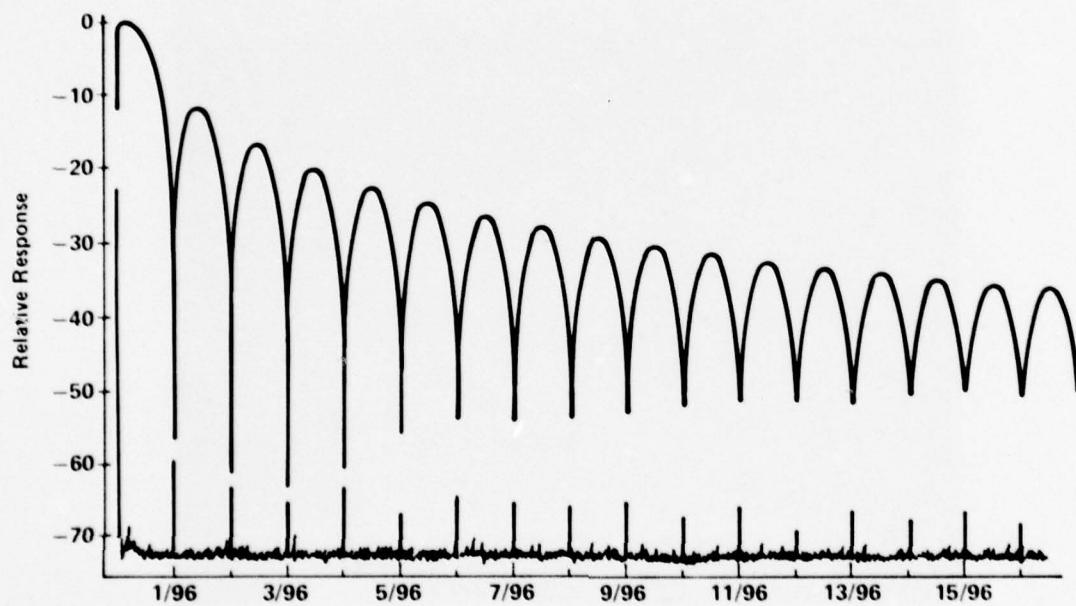
Fig. 10 Fixed pattern noise for a rectangular window (all unity weighting coefficients) realized with a 96 tap structure.

DISCUSSION

L.G.Gregoris

(1) How many capacitance bits is achievable practically with this organization?

(2) How fast can these weighting capacitors be cycled?

Author's Reply

(1) Our existing chips employ 7 and 9 bits data words. The accuracy of the LSB capacitor is one limiting factor. The available chip area (per tap) places a restriction on the MSB capacitor size. For a given application the word size can be chosen within the limits.

(2) For an n tap filter, n clock cycles are required to load the filter. In this case, somewhat more time is required because of the need to synchronize the filter and the microprocessor.

P.Carr

In communications applications, where you are charging your code cyclically, do you have a second channel where you store the new code?

Author's Reply

It is behaving in the same way as a normal transversal filter. So it takes you one circulation around the filter to get anything out. You have to wait at least one device length before charging your code.

DESIGN AND PERFORMANCE OF SAW-RESONATORS AND RESONATOR-FILTERS

W.H.Haydl, P.Hiesinger
Institut für Angewandte Festkörperphysik der Fraunhofer Gesellschaft
Eckerstrasse 4, D-7800 Freiburg, Federal Republic of Germany
and
G.Kohlbacher, P.Schmitt
AEG-Telefunken
Elisabethenstrasse 3, D-7900 Ulm, Federal Republic of Germany

SUMMARY

We have developed one-, two-, and three-port SAW-resonators on ST-quartz and YZ-lithium niobate employing aluminum metallization. Investigations have been carried out on the important parameters for the design of SAW-resonators, such as the reflection coefficient, number of reflector elements, reflector and transducer placement, and terminating impedance. Possible applications of SAW-resonators are as frequency stabilizing elements for oscillators and as narrow-band-pass filters.

1. INTRODUCTION

A new type of SAW-device, the SAW-resonator, which utilizes distributed reflectors to obtain efficient reflection of surface waves, was introduced recently (see Ref. 1). These devices are well suited for narrowband filters and oscillators in the VHF, UHF, and lower microwave range. The distributed reflectors form basically a Fabry-Perot cavity. Therefore SAW-resonators have several advantages over conventional SAW-devices: (1) smaller bandwidth (higher Q); (2) smaller size; (3) lower insertion loss.
For the past two years we have investigated the physical principles of these devices, and have developed them on temperature-stable, as well as temperature-senitive substrates (see Ref. 2). In order to keep the manufacturing process simple and thus minimizing costs, we have concentrated on realizing resonator structures by a simple aluminum metallization pattern produced by optical and electron-beam lithography on quartz and lithium niobate substrates. Because SAW-resonators are very phase-sensitive devices, their design is difficult. Out of many possible different designs we shall describe the most promising ones in some detail: control of the center frequency, permanent trimming to the final frequency by sputtering of thin films and in situ measurement, electronic tuning by means of varactors, low loss, series operation of several devices giving up to 100 dB sidelobe signal suppression, multimode operation, and oscillators with single- and multiple-mode resonators, respectively.

2. SAW-RESONATORS

A typical two-port device is shown in Fig. 1. The reflectors are periodic arrays of metal strips or grooves, $\lambda/4$ wide, with a center to center spacing of $\lambda/2$, where λ is the acoustic wavelength. Each reflector element reflects only a fraction of the incident wave, such that many hundred elements are necessary. Interdigital transducers are used to couple the signal to the cavity. The number of transducer electrodes used depends on the desired impedance level. The effective cavity length is greater than the spacing between reflectors and is many wavelengths, since the waves penetrate into the reflector. Since for most applications only one mode is wanted, the cavity length must remain small.
A typical frequency characteristic of a resonator is illustrated in Fig. 2. Reflectors are very narrowband components, typically of the order of 1%. The resonance lies within this very narrow region. Direct transmission between transducers is noticable outside this band at the -40 dB level. SAW-resonator Q's are in the range from 5000 - 50 000, and insertion losses as little as a few dB are possible.
The reflectors are periodic perturbations for surface waves. They should reflect anywhere from 0.5 to 2 % of the incident wave, and should be lossless, that is for example, surface waves should not be converted into bulk waves, or be lost in the form of electrical energy. Such effects will of course lead to lower resonator Q's. The reflections from the two edges of the perturbations are 180 degrees out of phase. By making the perturbations $\lambda/4$ wide, and spacing them with a period of $\lambda/2$, all reflections are in phase. Which reflector one chooses, depends often on the type of technology available.
Besides choosing between many kinds of reflectors, another variable is the placement of the transducers. In a resonator, acoustic energy is decreasing exponentially away from the center of the cavity. Placing the transducers outside the cavity leads to high Q's as well as high transducer impedances, because of the very loose coupling. Placing the coupling transducer inside the cavity, reduces its impedance highly. One- or two-port configurations are possible. Strongly piezoelectric materials such as lithium niobate require few finger pairs; quartz requires many, because it is much less piezoelectric.

Imporant parameters of a two-port resonator are the reflector-reflector spacing l_{rr}, and the transducer-reflector spacing l_{rt} as illustrated in Fig. 3. The presence of transducers inside the cavity makes the analysis and design of resonators very difficult. A resonance occurs when the total phase shift consisting of a propagation and a reflector part per roundtrip is a multiple of 2π. In order to determine the correct value for l_{rr}, one must know the relative difference in the propagation velocity in the various parts, such as under the transducers, reflectors, on the free surface, on metallized surfaces etc. In order to place the resonance onto the center frequency of the reflector stop-band, l_{rr} must be calculated from velocity data, and is generally not a multiple of d_r, the reflector periodicity. This cavity length correction is difficult to determine since no accurate velocity data is available. Another difficulty lies in the fact that transducers often are reflecting (an exception are double- or split-finger

transducers) such that the actual reflector includes the transducer. There are several possible positions for the transducers such that the total reflection can consist of two out-of-phase components, which in turn requires a correction in l_{rr}. If this correction is not made, a shift in the resonance frequency results.

In a resonator two standing wave patterns are possible as shown in Fig. 4. We assume here that the transducer does not disturb the reflection. As in the case of a transmission line with a short or open circuit in place of the reflector, a standing wave minimum or maximum can occur at the edge of the first reflector element, depending on whether the acoustic impedance step at that point is positive or negative. This impedance change ΔZ, in particular its sign, is generally determined more easily by experiment rather than by theory. It is not necessary that the transducer electrodes are placed at the standing wave maximum as shown. With a reduction in coupling, they may be placed off the maximum.

In Fig. 5 we illustrate a resonator design for YZ-lithium niobate with double-electrode transducers on the standing wave maximum. The resonator is an aluminum pattern on the lithium niobate substrate. The transducers are designed to have an impedance of 50 ohms at resonance. Because of the strong piezoelectric coupling, few electrodes are required. The impedance off resonance is much higher, several 1000 ohms.

For ST-quartz, more electrodes are required in order to couple to 50 ohms. For design purposes we illustrate the transducer impedance and insertion loss in Fig. 6 and Fig. 7. The resonator impedance is reduced by a factor of 10^2 to 10^3 (see Ref. 3), when the transducers are placed in the cavity of the resonator. Figure 8 illustrates a resonator with the transducer off the maximum. It has the advantage that the reflective properties and the velocity along the entire structure are constant, resulting in no out-of-phase reflection components, and with the frequency occuring exactly at the center of the reflector stopband.

Because the transducer is at a fixed distance from the reflector, and because the reflecttor has a certain phase and amplitude characteristic, the standing wave pattern will shift with frequency. When the transducers are at the standing wave maximum, the response is symmetric, otherwise we observe nulls when they are $\lambda/8$ off the maximum. The calculated amplitude and phase of a simple reflector are shown dotted in Fig. 9. We define the reflector stopband as the center lobe. High-Q resonance is possible where the reflection is near unity. If the transducers are reflecting, and placed on the standing wave maximum, the calculated response is shown by the solid curve. The strong phase distortion accounts for the observed frequency shift of the main resonance, and the appearance of other modes, as shown in Fig. 10. A lithium niobate-aluminum resonator response is shown in Fig. 11 with transducers on the standing wave maximum, and negligible reflection from the transducers. Amplitude and phase are symmetrical. A summary of transducer-reflector relation for various types of reflectors, for both single- and double-electrode transducers, is shown in Fig. 12. With its help one can analyze or predict the frequency response of resonators.

The frequency of a resonator may be varied by evaporating thin oxide films such as ZnO or Al_2O_3 onto the surface. This will cause a shift of the entire resonator characteristic. Another way of trimming the frequency is possible by changing the phase in the space between the reflectors only. In this case the reflector stopband is not affected and only the resonance peak will shift. For example, in materials with a very high piezoelectric coupling coefficient, the propagation phase between the two reflectors can be changed electrically. It is known that the velocity under a transducer depends on its terminating impedance. Thus by placing a third transducer between the reflectors, it is possible to observe a change in propagation phase. This phase change will be compensated by an equal but opposite change of the reflector phase, which is only possible, if the resonance frequency changes within the stopband of the reflector. Experimental results are shown in Fig. 13 for a transducer terminated in a capacitance, such as a varactor for instance. As the capacitance is increased, the transducer is increasingly short circuited, causing the frequency to decrease. The effect is only practical for materials such as lithium niobate, since the velocity change is proportional to the strength of the piezoelectric coupling.

In practice, if one designs a resonator, the probability is very great that the resonance does not occur at the stopband center. The reason for this is found in the phase-frequency relation per roundtrip. For a typical case we have plotted the position of the resonance versus frequency for small reflector spacings in Fig. 14. This case applies also when the transducer is part of the reflector and therefore placed off the standing wave maximum. If the spacing is slightly off n times $\lambda_f/2$, where λ_f is the acoustic wavelength in the space between the transducers, which it generally is, two modes are observed. For larger cavities, as illustrated in Fig. 15, more modes appear. The modes are not equally spaced, except near the center of the reflector stopband, as illustrated in Fig. 16. The mode spacing is about $v/2l_c$, where l_c includes the penetration depth into the reflectors.

We shall very briefly discuss a few more parameters. One of them is the width of the reflector stopband. It is of the order of 1 %, and smaller for smaller ΔZ, that is shallower grooves or thinner aluminum metallization. An approximate relation as well as experimental results are illustrated in Fig. 17. Another parameter is the number of reflector elements. It cannot be increased indefinitely in order to obtain as close to 100 percent reflection as possible. Because various losses exist for a propagating surface wave, experimentally one finds that the Q only increases up to a certain number of elements. This critical number is about $4/\Delta Z$, as shown in Fig. 18.

Acoustic energy is confined inside the structure of the resonator. This gives rise to transverse modes which appear as additional small resonance peaks near and above the main resonance peak. These transverse modes can be highly suppressed by exciting the waves mainly at the center of the structure by means of apodized transducers, as illustrated in Fig. 6.

So far we have only discussed simple resonators. Clearly, a variety of designs is possible. One class are resonators coupled either acoustically, or electrically as illustrated in Fig. 19. In (a) and (b) acoustic energy leaks into the next resonator. In (c) and (d) transducers are coupled. As an example we show case (c) in Fig. 20, where three resonators have been connected in series, resulting in 80 dB suppression off resonance.

Applications for resonators or resonator-filters are as low loss narrowband filters in the 100 - 1000 MHz range, as the frequency controlling element in oscillators either at VHF, UHF or microwave frequencies as illustrated in Fig. 21. SAW-devices are more robust and shock resistant than volume wave resonators. Frequency trimming and ageing problems are present, they are being investigated by a number of groups and we have great confidence that many more very useful devices and applications will emerge within the next few years.

3. DISCUSSION

Out of the many possibilities to realize a SAW resonator, two materials as well as two technologies have emerged. YZ-lithium niobate and ST-quartz are the two materials most often used. ST-quartz is the preferred material when temperature stability is important. Transducers are generally made of aluminum because of mass loading effects and economy. The reflectors may either consist of aluminum strips or etched grooves(ion, plasma, reverse sputtered, chemically).The SAW resonance phenomena may be used to realize either (1) narrowband(0.01-0.1% BW) and low insertion loss(0-6 dB) filters, (2) medium Q (1000-10000) resonators, and (3) high Q (10-100 000) resonators. We have described here the general principles of simple SAW resonators and filters. Medium Q devices can already be obtained with simple technology, such as an aluminum structure for transducers and reflectors. With increasing frequency however, we have observed that the losses (believed to be propagation and scattering losses) are becoming excessive, limiting the Q of the devices. For example, the resonator of Fig. 10 has a Q of 12 000 at 30 MHz, 8000 at 150 MHz, and 4000 at 300 MHz. Our highest Q's (unloaded) with aluminum structures have been 20 000 at 30 MHz.

For high Q devices, the number of transducers inside the cavity, as well as the number of transducer electrodes, should be kept to a minimum, which in turn means loose coupling and high impedance levels. Reflectors should be of the groove type. By employing self-aligning photolithographic techniques and plasma etching, the technology may also be kept simple.

As an example, for the design of a simple resonator , we recommend the following guidelines:

a.)Choose ΔZ to be between 0.5 and 2%. The product of ΔZ and the number of reflector grooves must be at least 4. For high Q select a low ΔZ . $\Delta Z = 0.8$ h/λ , where h is the groove depth and λ is the acoustic wavelength.

b.)Use double-electrode transducers, operating either at the fundamental or the third harmonic frequency, placed at the standing wave maximum.

c.)Use the largest possible reflector spacing for single mode operation. Since the relative velocity differences within the cavity (reflector, transducer, free surface) are not known with sufficient accuracy, it is generally not possible to design the reflector separation l_{rr} (Fig. 3.4) correctly such that $\phi_{rr} = n\pi$. A large l_{rr} makes the resonance frequency less sensitive to deviations of ϕ_{rr} from $n\pi$. It also reduces the electric feedthrough in the case of two-port devices. The maximum l_{rr} can be estimated from the relation $(l_{rr}/\lambda)_{max} = (2\Delta f_b /f_0)^{-1}$. Using the relation between $\Delta f_b/f_0$ and ΔZ in Fig. 17 we find that $(l_{rr}/\lambda)_{max}$ =120 for a ΔZ of 0.5%, 80 for 1%, and 40 for 2%.

d.) The resonance frequency may be shifted within the reflector stop band bei changing the propagation or reflector phase. We have done this by depositing Al_2O_3 in the space between reflectors, thus increasing the velocity there and hence decreasing the propagation phase. An increase in the resonance frequency is thus obtained. However, disturbing the cavity with metal- or dielectric films will have adverse effects on the Q , and probably also on the long-term stability.

The approach we recommend is to calculate the necessary phase change. Knowing the phase slope of the reflector$(d\phi/d(f/f_0)=-\pi/\Delta Z)$ as well as of l_{rr}, it is possible to obtain the necessary Δl_{rr} for the new design. Because of ageing problems it is difficult to place the operating frequency of a resonator at an exact prescribed value. Ageing is being investigated by a nuber of groups at the present time. Ageing effects of the substrates have been observed by us as well. They are as high as 10^{-4} over a period of one year. Substrate ageing may be reduced by an accellerated ageing treatment of the substrate. Velocity differences within the substrates have been studied by us, and are several times 10^{-4} in quartz, and 10^{-2} in lithium niobate. Because lithium niobate is very temperature sensitive, $(10^{-4}/^{\circ}C)$, high Q resonators may be temperature tunable.

REFERENCES

1. See 1974 through 1977 IEEE Ultrasonics Symposium Proceedings for many papers on SAW-resonators.

2. W.H.Haydl, B.Dischler, P.Hiesinger, "Multimode SAW Resonators - A Method To Study The Optimum Resonator Design", 1976 Ultrasonics Smposium Proceedings, IEEE Cat.# 76 CH 1120-5SU, p. 287-296.

3. E.J.Staples, J.S.Schoenwald, R.C.Rosenfeld, C.S. Hartmann, "UHF Surface Acoustic Wave Resonators", 1974 Ultrasonics Symposium Proceedings, IEEE Cat.#74 CHO 896-1SU, p. 245-252.
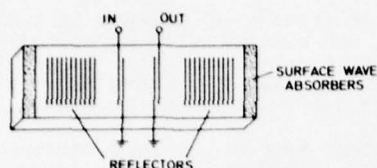
Fig.1:Schematic representation of a two-port resonator.
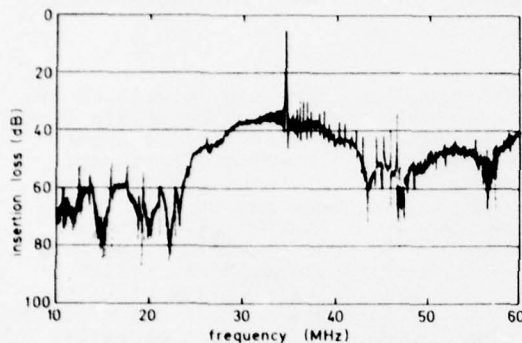
Fig.2:Typical frequency characteristic of a two-port saw resonator filter on YZ-lithium niobate. The sharp peaks throughout the spectrum are plate modes.
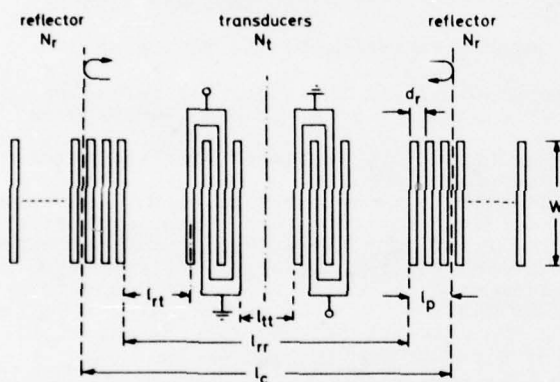
Fig.3:Two-port SAW resonator parameters.
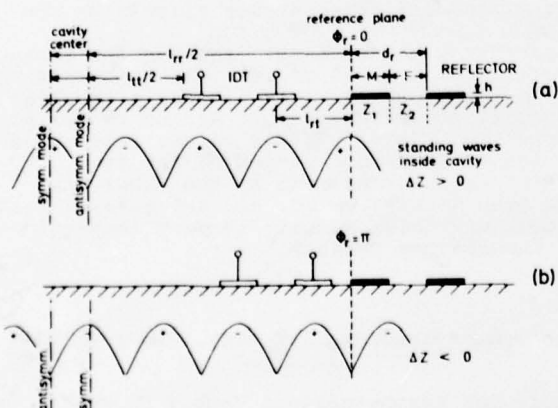
Fig.4:Longitudinal cross-section of a resonator. Interdigital transducers(IDT) are shown positioned on standing wave maximum.
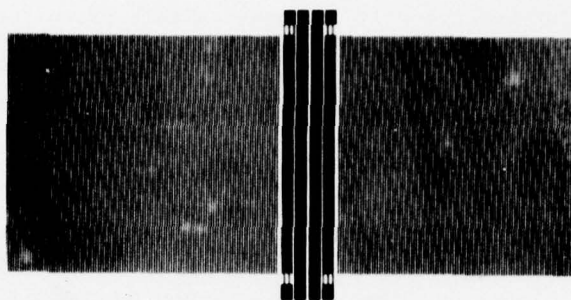
Fig.5:Actual layout of a two-port SAW resonator on YZ-lithium niobate with double-electrode transducers having 2.5 finger pairs and two rf shields. Only a small portion of the reflecting strips are shown.
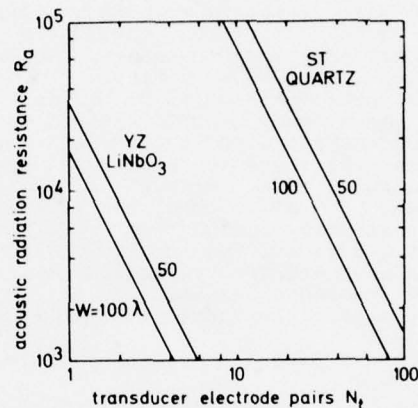
Fig.6:Calculated acoustic radiation resistance for single electrode transducers.
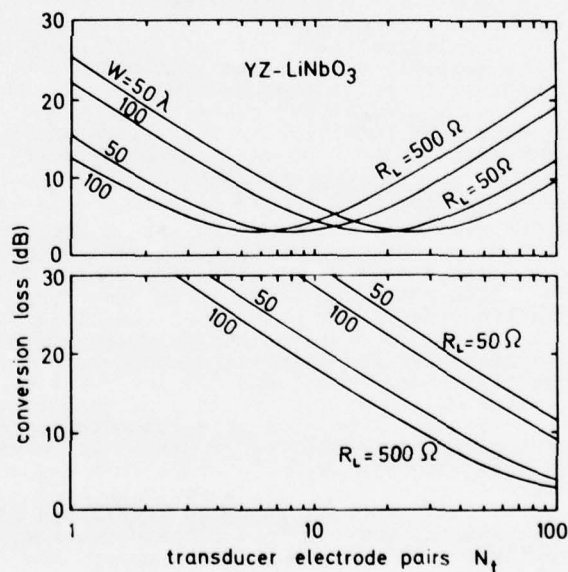
Fig.7:Calculated conversion loss for single electrode transducers.

Fig.8:Two-port SAW resonator with cos x weighted transducers on ST quartz.
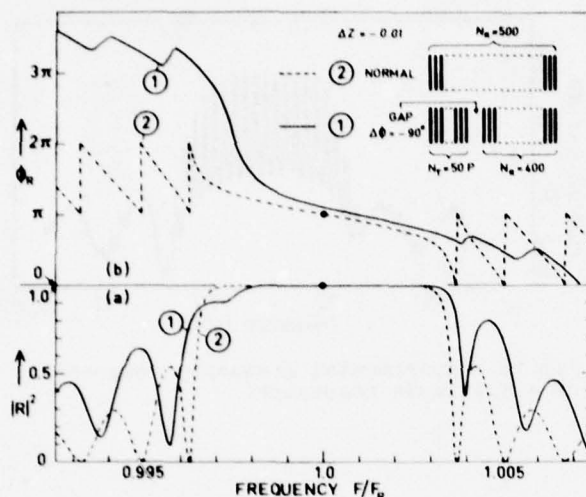
Fig.9: Calculated distortion of reflection phase and amplitude caused by the phase difference between reflector and single-electrode transducer. Curves 1 are for transducer electrodes placed on the standing wave maxima; curves 2 are for a single reflector with the same total number of reflecting elements and the same ΔZ. (a) power reflection coefficient, (b) reflector phase at reference plane.
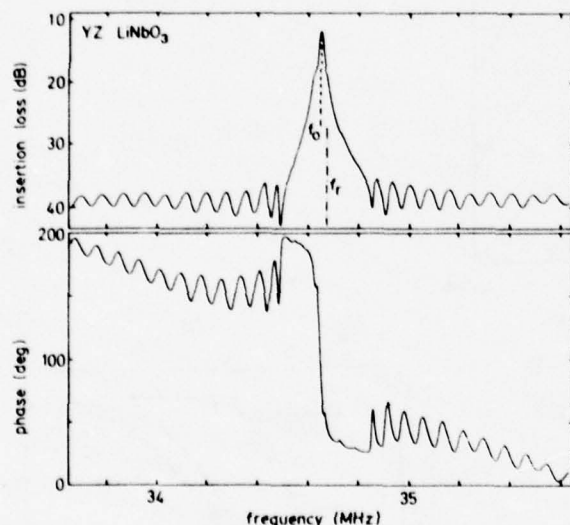
Fig.11: Detailed amplitude and phase characteristics of a resonator on YZ-lithium niobate.
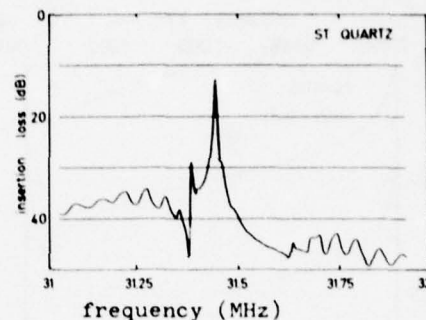
Fig.10: Experimental frequency response for a two-port resonator with 30 single-electrode pairs on the standing wave maxima, with the resonance peak shifted to lower frequencies.

$\Delta Z > 0$ $(\Phi_r = 0)$ open Al-YZ LiNbO$_3$, grooves
$\Delta Z < 0$ $(\Phi_r = \pi)$ shorted Al-YZ LiNbO$_3$, Al-ST quartz, ridges

Fig.12: Summary of transducer-reflector interactions in SAW resonators showing schematically the frequency responses over the reflector stop band. The solid lines are resonance peaks.
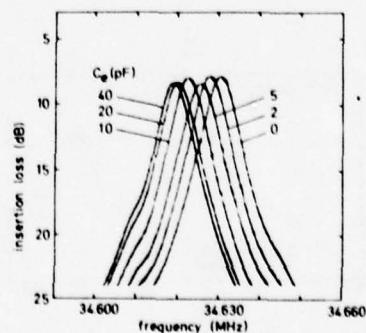
Fig.13: Tuning characteristics with an external capacitance for three-port SAW resonator on YZ-lithium niobate.
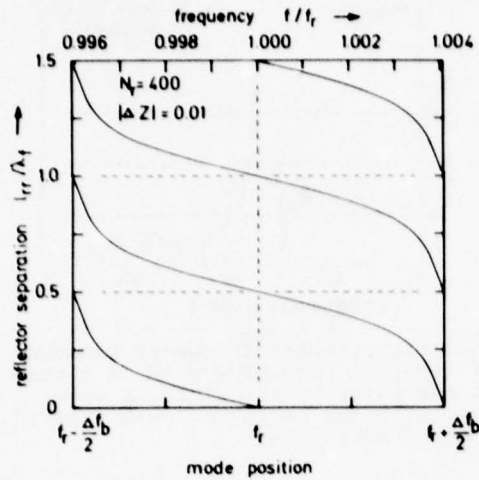
Fig.14: Calculated positions of the resonance modes for small reflector spacing.
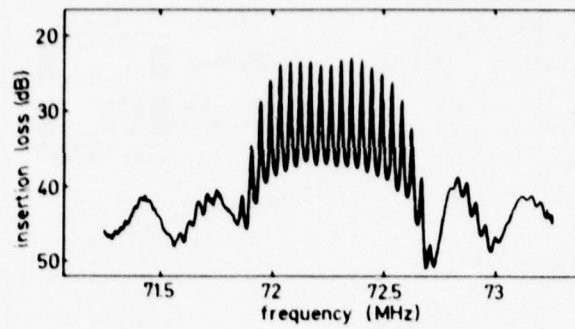
Fig.16: Experimental frequency response of a multimode resonator.
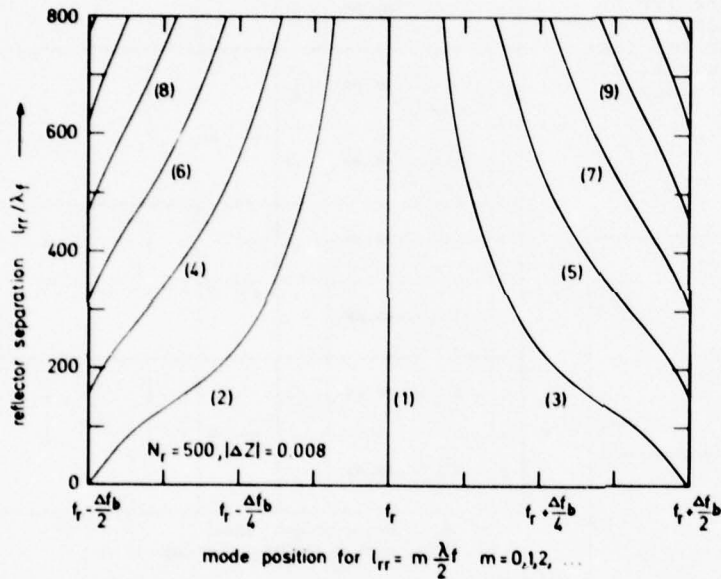
Fig.15: Calculated positions of the resonance modes.
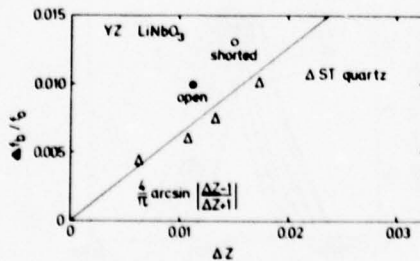
Fig.17: Experimental and theoretical values of the reflector base frequency Δf_b for $N_r > N_{rc}$ (see Fig.18).
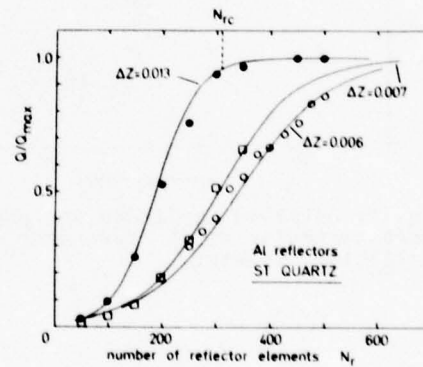
Fig.18: Measured (dots) and calculated(lines) variation of normalized Q.

coupled resonators

(a)

(b)

(c)

(d)
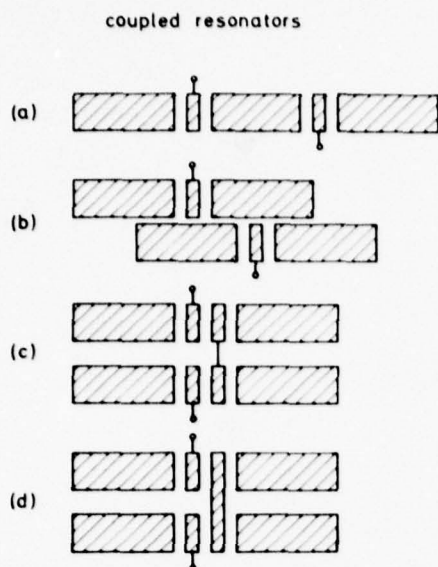
Fig.19: Representative structures of several coupled resonators.

SAW RESONATOR FEEDBACK OSCILLATOR
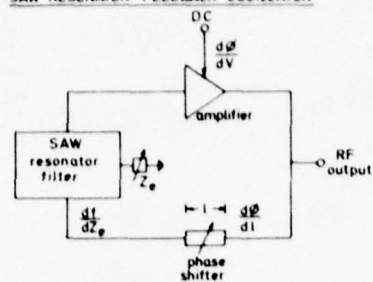
SAW - XTAL CONTROLLED MICROWAVE SOURCE

Fig.21: SAW oscillators.
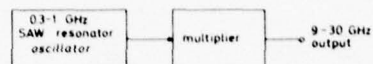
Fig.20: Frequency response of three series connected resonators.

SYSTEMS APPLICATIONS OF SAW FILTERS AND DELAY LINES

Paul H. Carr
Deputy for Electronic Technology (RADC/AFSC)
Hanscom AFB, Massachusetts 01731

SUMMARY

This paper reviews the applications that SAW (Surface Acoustic Wave) filters and delay lines have to communications and radar systems. The development of SAW resonators at frequencies up to 1 GHz is reported together with their promise for stabilizing precision oscillators. Such oscillators could be useful, for example, as clocks for CCDs. Operation directly at 1 GHz eliminates the multiplier chains required for conventional quartz oscillators. The use of miniature SAW transversal filters for banks of contiguous filters is shown to be important in reducing the size of fast frequency synthesizers and multiplexers for spread-spectrum communications applications. The design of a 9.1 μsec delay line with all spurious echoes 70 dB down is described together with its radar signal processing application.

1. INTRODUCTION

The signal processing capabilities of SAW and CCD are generally complementary. SAW (Surface Acoustic Wave) devices are capable of bandwidths up to 1 GHz but are limited in time delay to 100 μsec, while CCD (Charge Coupled Devices) can have time delays up to 0.1 sec but bandwidths are generally tens of megahertz. (See Fig. 1.) The two technologies are also complementary in that CCD devices require a clock which can be provided by SAW technology. This paper will focus on the very small time delay section of Fig. 1 and will be concerned with SAW bandpass filters and resonators. These components are much smaller than their electromagnetic counterparts and are lower cost, as they are fabricated by the same photolithographic techniques used for integrated circuits. We will describe applications of these filters towards stabilizing precision oscillators and reducing the size of fast frequency synthesizers and frequency multiplexers. Finally, the application of SAW delay lines to radar signal processing applications will be described including the techniques required to maintain all spurious responses 70 dB below the 9.1-μsec delay time.

2. SAW RESONATORS

Recently Laker, et. al. (LAKER, K.H..., 1977) have reported on the development of SAW resonators with Qs from 4000 to 5000 at 0.8 GHz. (Fig. 2) Such high-Q resonators are of importance in stabilizing the frequency of precision oscillators used in communication and radar systems. Figure 3 shows the two small transducers in the center which are used for the electromagnetic input and output. Acoustic standing waves are built up between the long reflectors consisting of 1-μm-wide aluminum lines on 2-μm centers. The high-Q of the acoustic resonance couples the input to the output transducer.

A simple estimate of the Q can be obtained by assuming a reflectivity parameter for each grating line and accounting for propagation and diffraction loss (BELL, D.T.,JR...1976). Because stripe reflection coefficient data are scarce at these high frequencies, the value of $r = -.006$ gave reasonable agreement with experiment. For this case, the unloaded Q-factor in air was calculated (LAKER, K.R...1977) to be 5090, in air and 5540 in vacuum. Figure 2 shows that the increase in Q obtained by evacuating the resonator was in good agreement with that calculated by simple theory. The Q due to propagation loss in vacuum on quartz is 12,500 at 0.8 GHz. This when combined with diffraction loss results in a Q of 8000, which can be approached as other loss mechanisms are minimized.

Figure 4 shows a comparison of SAW and conventional electromagnetic resonators. The helical resonator is one-half wavelength long. The dielectric resonator uses strontium titanate to reduce its size. The SAW resonators have smaller size, substantially higher Q, and the advantage of being planar. The last factor makes them compatible with integrated circuits and microstrip and contributes to low-cost mass production.

Figure 5 shows a comparison between a SAW stabilized resonator at 1 GHz and a conventional 20 MHz crystal oscillator with the multipliers, amplifiers, and filters necessary to produce an output at 1 GHz. Quartz crystal oscillators cannot operate at frequencies substantially higher than 20 MHz, as the thickness of the crystal must be either one-half acoustic wavelength thick or an odd multiple thereof. For frequencies above 20 MHz the quartz crystal becomes too fragile. The planar SAW device operating directly at 1 GHz eliminates the multiplier chain with an approximate savings in size and weight of a factor of 20 together with increased reliability. The SAW filter requires the use of only one surface of the crystal. The remaining five sides may be chosen, for example, to reduce the vibration sensitivity.

Figure 6 shows the insertion loss of the SAW resonators over a much broader range than in Figure 2. The insertion loss of the unmatched resonator is 33 dB (LAKER, K.R...,1977). By crude stub tuner matching the insertion loss was reduced to 22 dB and the Q-factor degraded to about 3900. The lowest insertion loss of 10 dB was obtained by inductive matching of the input and output transducers.

The off-resonance coupling between the input and output is due to electromagnetic leakage and acoustic cross talk. Fildes and Hunsinger (FILDES, R.D......1976) have shown that unidirectional transducers can be used to substantially reduce the acoustic cross-talk between the transducers.

The long term stability or aging of SAW resonators is not as good as conventional resonators. Adams and Kusters (ADAMS, CA.,....1977) have observed long term stabilities of no better than 1 part in 10^8 per day as compared to bulk-wave resonators, whose long term aging is better than 1 part in 10^{10} per day. As the relatively young SAW technology matures, the aging should also improve.

3. FAST FREQUENCY SYNTHESIZER

We shall now discuss important applications of SAW bandpass, transversal filters, which in analogy to resonators are much more compact than their electromagnetic counterparts. SAW transversal filters below 100 MHz are being mass produced at a cost of under $2.00 for use in color TV and games. This paper will

discuss SAW filters at higher frequencies, where they are even smaller. A 21 channel contiguous filter bank from 500 to 600 MHz occupies an area of only 9 by 25mm. (See line drawing in Figure 7 and photograph in Figure 8.) Frequencies can be switched by the PIN diodes in less than 5 nsec, a function useful for coherent frequency-hop coding in spread spectrum communications systems (LAKER, K.R.,....1976).

The desired frequency synthesis is achieved by feeding a uniform comb spectrum, which is harmonically related to a stable reference oscillator or clock, into the filterbank. Each SAW filter is tuned to one of the different frequencies of the comb spectrum. Hence, a given frequency is continuously available at the output of the filter. The spectral purity of this signal is as good as the out-of-band rejection of the filter. Out-of-band rejections in excess of 60 dB are achievable. The stability of the signal is the same as that of the reference clock. The switching speed is that of the diode switch, whose on-off-ratio can be held in excess of 60 dB.

Figure 9 shows a mixer technique of using two of the above basic building blocks to substantially increase the number of available frequencies (SLOBODNIK, A.J.,JR.,....1976). As illustrated, one tone from the 9 channel filter bank is mixed with one from the 7 channel bank to obtain 63 tones with only 16 filters. The frequencies of each of the banks is carefully chosen so that, for a given mixing stage, only the desired sum or difference frequency falls within the bandwidth of the broadband electromagnetic filter following the mixer (thereby eliminating mixer spurious). The second mixer stage can be used to produce the desired output frequencies. Figure 9 also illustrates the multipliers and snap-diode comb generators which can be used to implement the dual comb generator.

Frequencies in the low gigahertz range can be met with sub-micrometer linewidth interdigital transducers using either electron-beam fabrication or direct optical projection from 10X masters, (BUDREAU, A.J.... 1975). To reduce the demands on the device fabrication, these high operating frequencies may also be achieved by operating filters at third overtone modes or by using high velocity substrate materials such as aluminum nitride-on-sapphire thin films and beryllium oxide, whose velocities are of the order of 6000 m/sec, (roughly twice that of ordinary materials such as quartz), (BUDREAU, A.J.,....1975).

SAW filters with 0.7 μm linewidths and spaces on lithium tantalate and with a frequency of 1.03 GHz were fabricated using direct optical projection (BUDREAU, A.J.....1975). Advantages of this technique are (1) 10X photolithographic masters are less expensive than one-to-one masters, (2) the master need not be in physical contact with the substrate thereby increasing its lifetime, and (3) ease of varying the magnification slightly about 10X for fine frequency tuning. The final one-to-one image is obtained with a high resolution reducing lens directly in photoresist on the substrate. Figure 10 illustrates high quality 0.6 and 0.9 μm wide metal transducer lines made by direct optical projection (KEARNS, W.J.,....1977).

4.   MULTIPLEXERS

Contiguous banks of SAW filters can be used, as shown in Figure 11, to make compact frequency multiplexers. The single electromagnetic input in Figure 11 comes through the 3 dB input attenuation (whose purpose is to keep the input voltage-standing-wave-ratio below 2) to the 8 input transducers. These are connected electrically in parallel to keep the triple-transit spurious below 45 dB (SLOBODNIK, A.J.,JR.,....1975). This novel technique eliminates the need for an electronic power-divider at the input, with its associated size and losses. The 8 output transducers, which are barely visible in the photograph, are connected to 8 matching output inductors. These, in turn, are connected to the 8 coaxial cables going to the output connectors on the front of the package. It is apparent that these are the largest size determining elements in this laboratory prototype package. These connectors would not, of course, be required in a systems application, where the filterbank would be directly connected to the output electronics. Note also that there are actually 16 SAW filters visible on the two lithium tantalate chips. The extra filters would not be necessary due to the very high yields recently achieved with direct optical projection (KEARNS, W.J.,...1977).

A photograph of the insertion loss (10 dB per division) versus the frequency (20 MHz per division) for the 8 output channels is shown in Figure 12. The insertion loss across the band was 33.5 dB plus or minum 1.5 dB. The insertion loss could have been at least 4 dB lower if there had not been a 45 dB triple-transit requirement. Van de Vaart and Solie (VAN DE VAART, H.,....1976) have obtained 10 dB insertion loss for an 8-channel multiplexer ranging from 190 to 232 MHz by the novel use of multistrip couplers for dividing the acoustic power on a lithium niobate substrate. The triple-transit was only 20 dB down, however.

The multiplexers under discussion are useful for (1) frequency division multiplexing to obtain optimum use of the bandwidth capacity of a communications channel, and (2) channelized receivers for the real-time spectral analysis of pulses. Filter responses useful for pulses are different from those for continuous waves (of the sort used for the frequency synthesizer). For pulses, filters are sought to be highly selective in the frequency domain and relatively distortionless in the time domain. One approach is to use contiguous Butterworth filters in a double-detection signal processing scheme (SLOBODNIK, A.J., JR....1975). In Figure 12, we have used a recently discovered (SLOBODNIK, A.J.,JR.,....1975) Flat Exponential Filter (FEF) function, which has a flat in-band response, and an exponential out-of-band response. In contrast to a Butterworth filter, which is straight-forward to implement with lumped constant elements, the Flat Exponential Filter, is straightfoward to implement with SAW transversal filters. When compared to a two-pole Butterworth filter, the FEF has better frequency selectivity, which is surprisingly not sacrificed in the time domain. This may well be the first of a series of new filter functions which are easier to implement with SAW rather than conventional filters.

5.   LOW SPURIOUS DELAY LINE

Finally, the application of a SAW delay line to a radar signal processing application will now be described. The most critical goal was to maintain the triple-transit and other spurious responses below 60 dB at the 30 MHz intermediate signal processing frequency (See Table I). At this frequency, bulk waves had in the past been difficult to suppress due to the fact that they have negligible attenuation. The solution to this problem was to use an orientation of lithium tantalate for which the piezoelectric coupling to volume and plate waves was very weak (CARR, P.H.,...1976). The MDC (Minimum Diffraction Cut) of lithium tantalate

discovered by Slobodnik (SLOBODNIK,A.J.,JR.....1975) to have diffraction spreading retarded with respect to an isotropic medium by a factor of 20, has this desirable property. Figure 13 illustrates that the longitudinal volume wave spurious, which arrives in about 5 µsec for YZ lithium niobate, is absent for the MDC lithium tantalate.

A line drawing of the delay line is shown in Figure 14. The transducers were separated by 3.10385 cm. The bottom of the substrate was rough ground and tapered to minimize any possible interference from bulk waves. The lithium tantalate slab was epoxied onto a stainless steel backing plate. This gave a strong, rigid structure, which made it an easy task to maintain flatness of the top surface to within three optical wavelengths. Randomly spaced holes were drilled into the stainless steel in order to scatter any possible volume waves. The electromagnetic leakage was maintained well below the 70 dB level by (1) grounding opposite sides of the transducer and, (2) by maintaining the surface of the crystal within 0.025mm of the bottom of the cover. This delay line illustrates that it is indeed possible to keep the spurious responses of SAW components below 70 dB.

6. CONCLUSION

We have reviewed the promise that SAW devices have for stabilizing precision oscillators directly at frequencies up to 1 GHz. Such oscillators are useful, for example, as clocks for CCDs. Operation directly at 1 GHz eliminates the bulky multiplier chains required with conventional quartz oscillators. We have shown how the small size of SAW transversal contiguous banks can play an important role in reducing the size of fast frequency synthesizers and multiplexers for important communications applications. Finally, we have shown how it is possible to design SAW delay lines for radar signal processing applications with all spurious echoes below 70 dB.

REFERENCES

1. ADAMS, C.A. AND KUSTERS, J.A., 1977. "Deeply Etched SAW Resonators," 31st Annual Frequency Control Symposium.

2. BELL, D.T., JR., AND LI, R.C.M., 1976, "Surface-Acoustic-Wave Resonators," Proc. IEEE, Vol 64, p 711.

3. BUDREAU, A.J., KEARNS, W.J., AND CARR, P.H., 1975. "State-of-the-Art in Microfabrication of SAW Devices," 1975 Ultrasonics Sympsoium Proceedings, IEEE Cat.No.75 CHO 994-4SU, p 458.

4. CARR, P. H., FENSTERMACHER, T.E., SILVA, J.H., KEARNS, W.J., AND STIGLITZ, M.R., 1976, "SAW Delay Line with All Spurious 70 dB Down," 1976 Ultrasonics Symposium Proceedings, IEEE Cat. No. CH1120-5SU, p459.

5. FILDES, R.D., AND HUNSINGER, W.J., 1976, "Application of Unidirectional Transducers to Resonator Cavities," 1976 Ultrasonics Symposium Proceedings, IEEE Cat. #76-CH1120-5SU, p. 303.

6. KEARNS, W.J., 1977. "Private Communication."

7. LAKER, K.R., BUDREAU, A.J., AND CARR, P.H., 1976, "A Circuit Approach to SAW Filterbanks for Frequency Synthesis," Proc. IEEE., Vol 64, p 692.

8. LAKER, K.R., SZABO, T.L., AND KEARNS, W.J., 1977, "High-Q-Factor SAW Resonators at 780 MHz," Electronics Letters, p. 97.

9. SLOBODNIK, A.J.,JR., FENSTERMACHER, T.E., KEARNS, W.J., ROBERTS, G.A., AND SILVA, J.H., 1975. "A Minimum Diffraction Lithium Tantalate Substrate for Contiguous SAW Butterworth Filters," 1975 Ultrasonics Symposium Proceedings, IEEE Cat. No. 75 CHO 994-4SU, p. 405.

10. SLOBODNIK, A.J., JR., BUDREAU, A.J., KEARNS, W.J., SZABO, T.L., AND ROBERTS, G.A., 1976, "SAW Filters for Frequency Synthesis Applications," 1976 Ultrasonics Symposium Proceedings, IEEE Cat. #76 CH1120-5SU.

11. SLOBODNIK, A.J.,JR., LAKER, K.R., FENSTERMACHER, T.E., 1975, "A SAW Filter with Improved Frequency and Time Domain Characteristics for the Frequency Measurement of Narrow RF Pulses," 1975 Ultrasonics Symposium Proceedings, IEEE Cat. #75, CHO 994-4SU, p. 327.

12. VAN DE VAART, H., AND SOLIE, L.P., 1976, "Surface-Acoustic-Wave Multiplexing Techniques," Proc. IEEE Vol. 64, p 688.
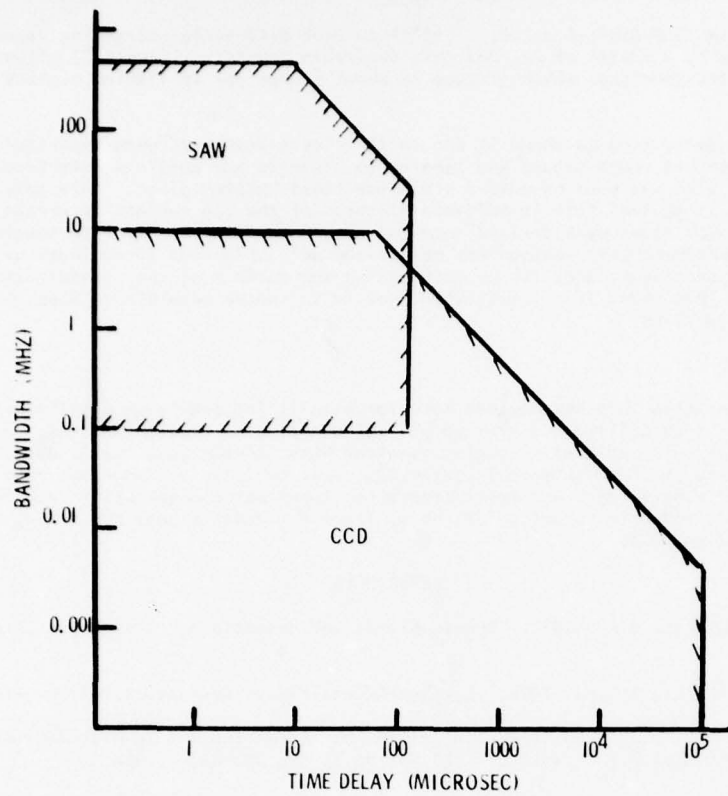
Fig. 1   Plot of the signal processing bandwidth versus processing time for Surface
Acoustic Wave and Charge Coupled Devices

ST-QUARTZ TWO-PORT RESONATOR

AIR
(Q = 4400)

VAC
(Q = 4960)

780.618 MHz
(50 KHz / DIV)
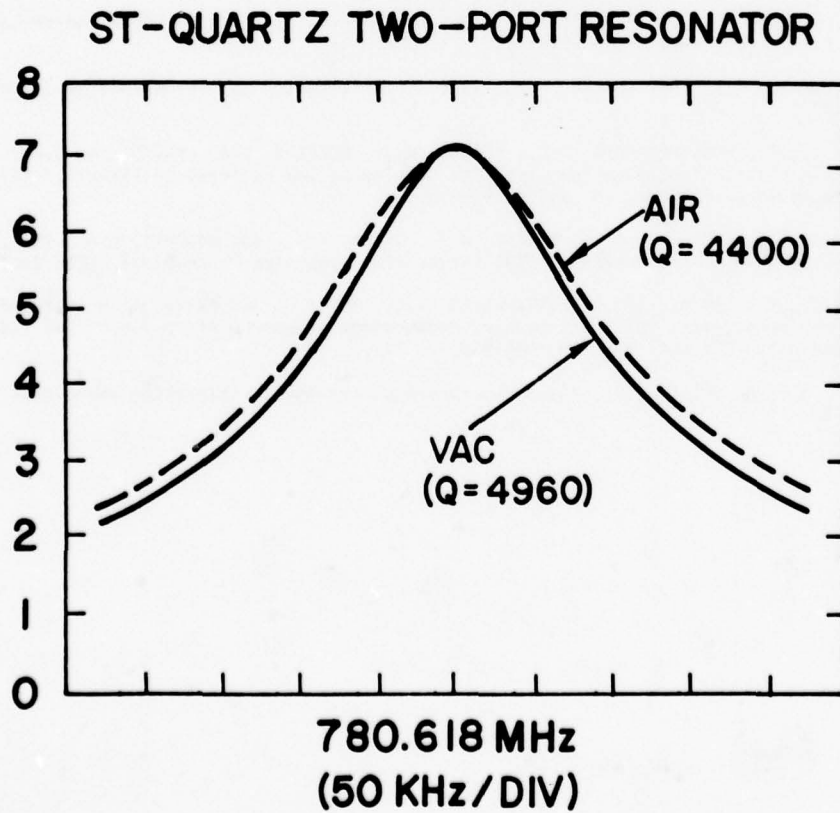
Fig. 2   Plot of the amplitude response versus frequency for 780 MHz SAW resonator in air
and in vacuum

TWO-PORT SAW RESONATOR

23.37 µm — 270.5 µm — 17.37 µm
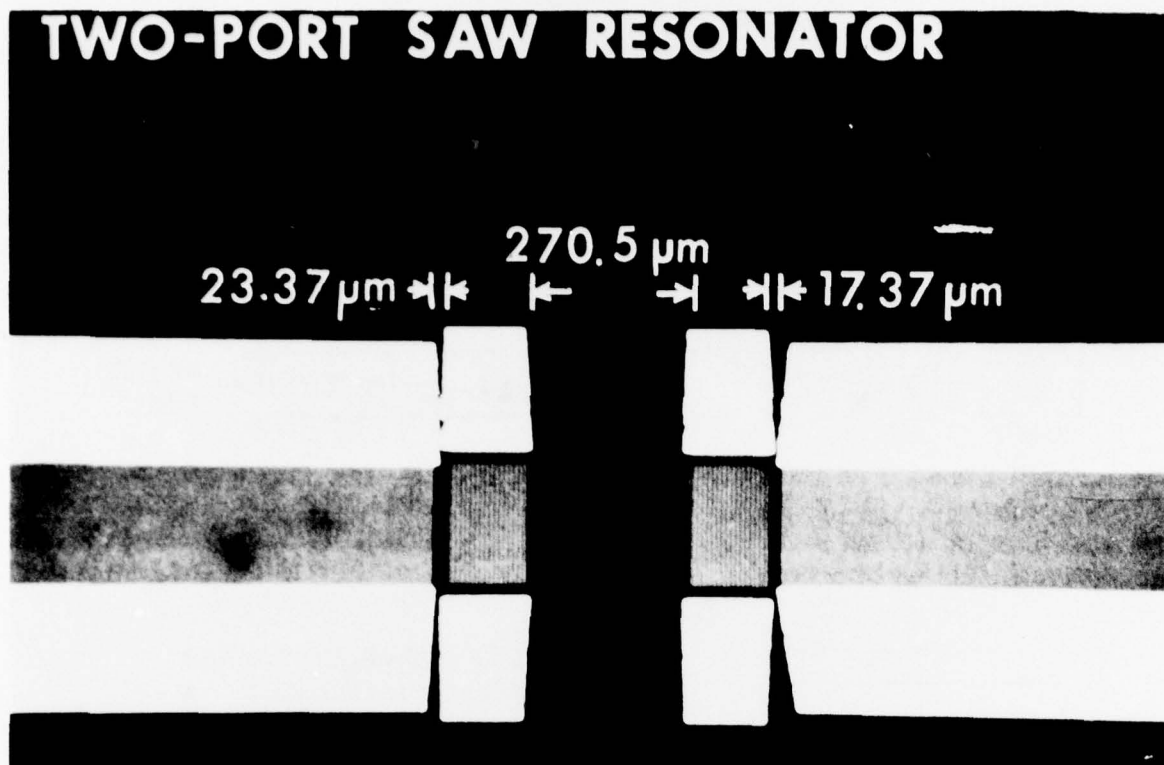
Fig.3   Photograph of a SAW resonator

Fig.4   Comparison of different resonators at 1 GHz. On the left, is a helical resonator (Q = 1000), in the center is a high dielectric constant resonator (Q = 1000), and on the right is a quartz chip with 5 SAW resonators with Q for 4000 to 5000

SIZE AND WEIGHT SAVINGS WITH SAW OSCILLATORS

CONVENTIONAL 1GHZ FREQUENCY SOURCE VOLUME ~100 CU. INCHES
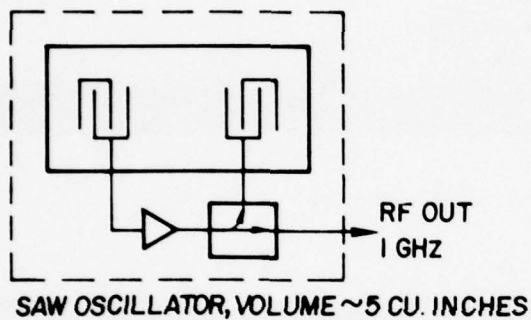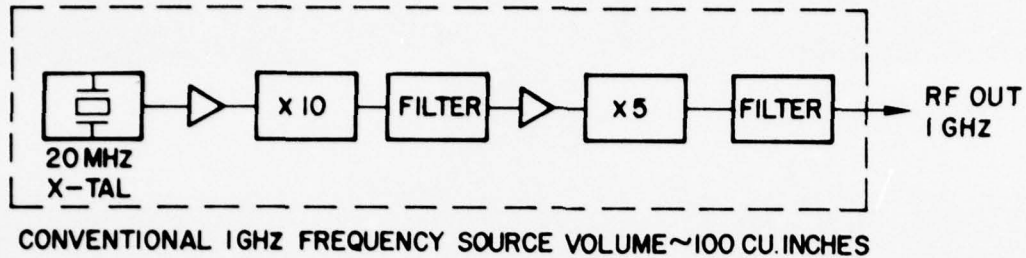
SIZE AND WEIGHT SAVINGS
APPROXIMATELY 20:1

SAW OSCILLATOR, VOLUME ~5 CU. INCHES

Fig. 5   SAW oscillators operate directly at 1 GHz and thereby eliminate the bulky multiplier
chains required with conventional quartz crystal oscillators

IL(dB)
(10dB/DIV.)

−30 —

UNMATCHED
Q=4398

780 5
FREQUENCY (MHz)
(10 MHz / DIV.)

IL(dB)
(10dB/DIV.)

−20 —

MATCHED
Q=3923

780.5
FREQUENC (MHz)
(10 MHz/DIV.)
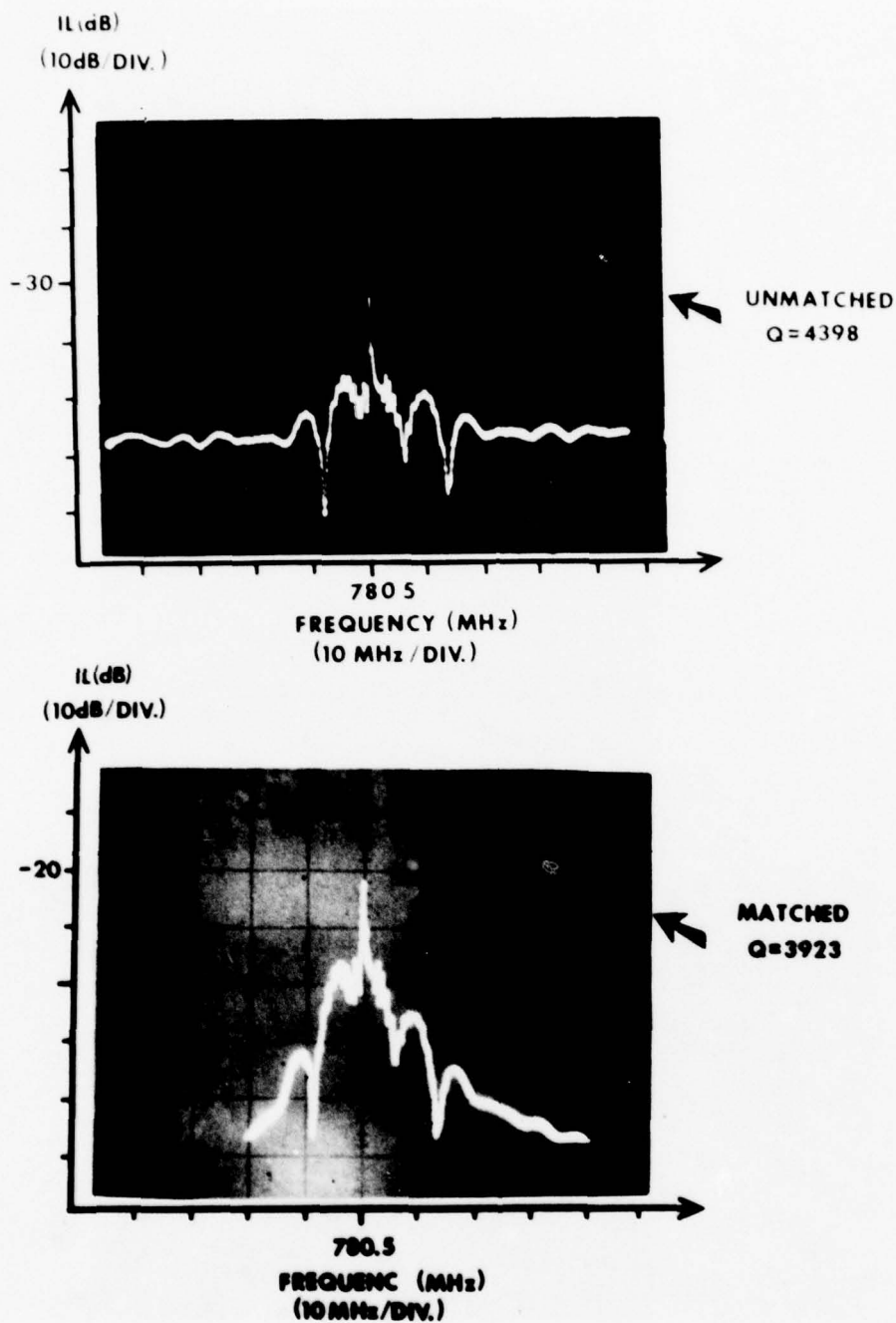
Fig.6 Insertion loss versus frequency plot for the resonator shown in Figures
2 and 3 for unmatched (top) and matching with stub tuners (bottom)
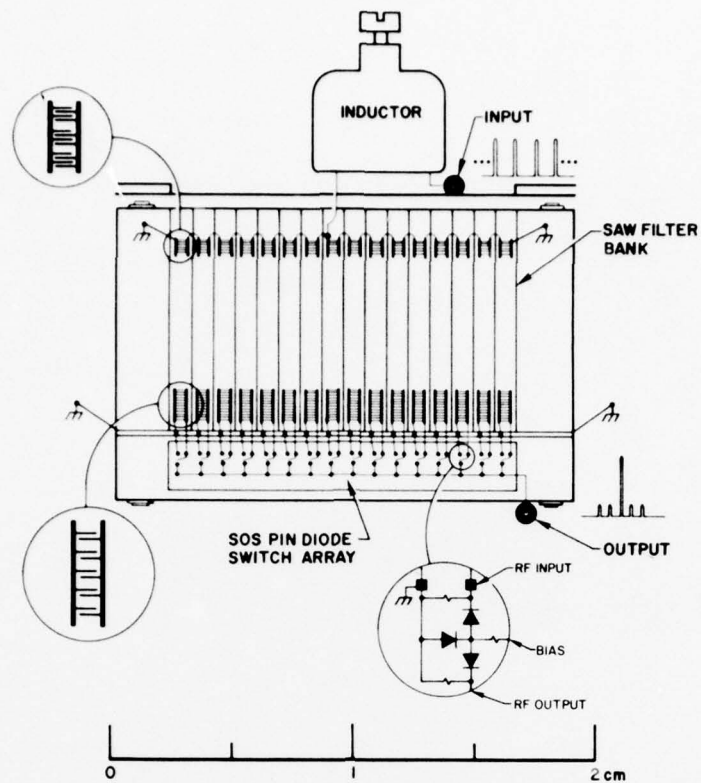
Fig.7 Drawing of UHF Frequency Synthesizer using a compact contiguous
SAW filter bank
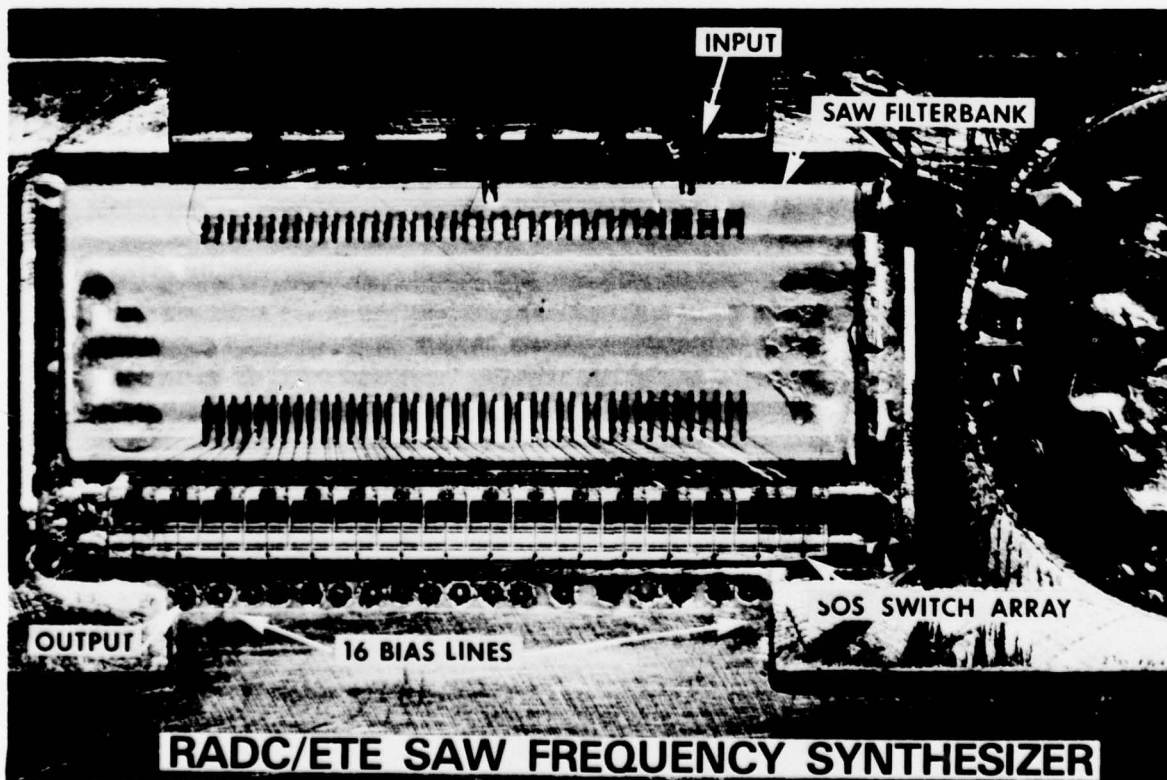
Fig.8 Photograph of the Frequency Synthesizer of Figure 7
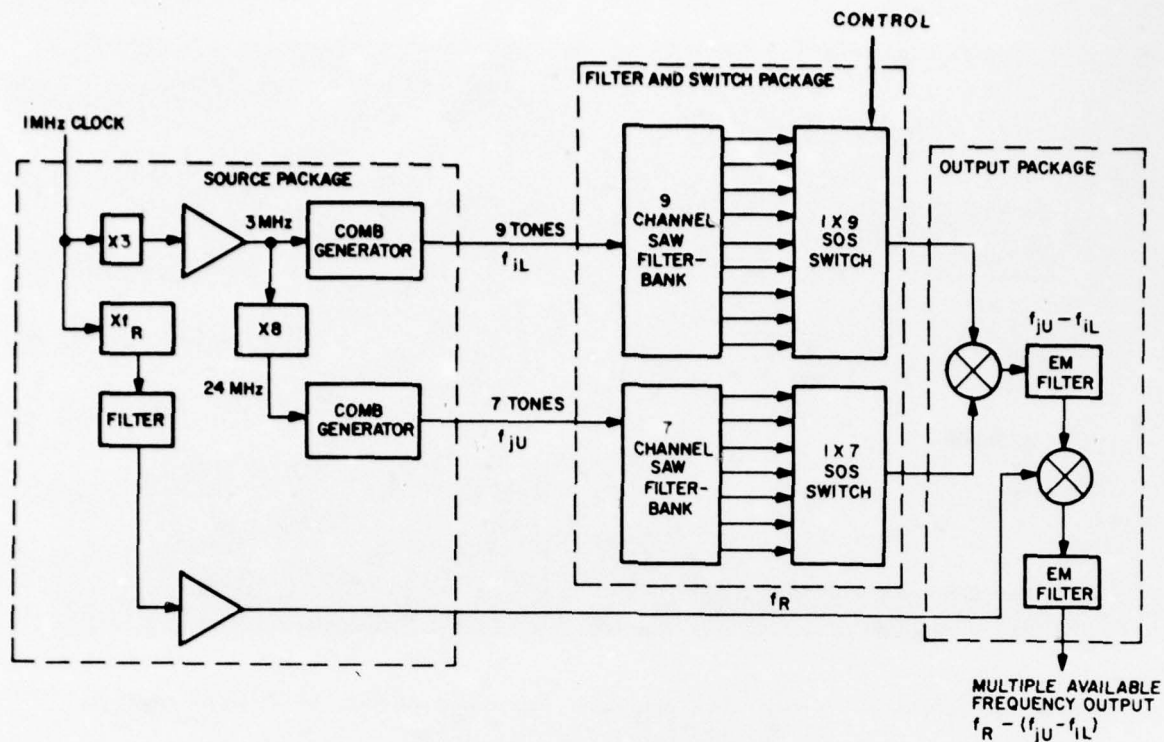
Fig.9 Schematic of how the outputs from two "building-block" frequency synthesizers of Figures 8 and 9 can be mixed to increase the number of available frequencies
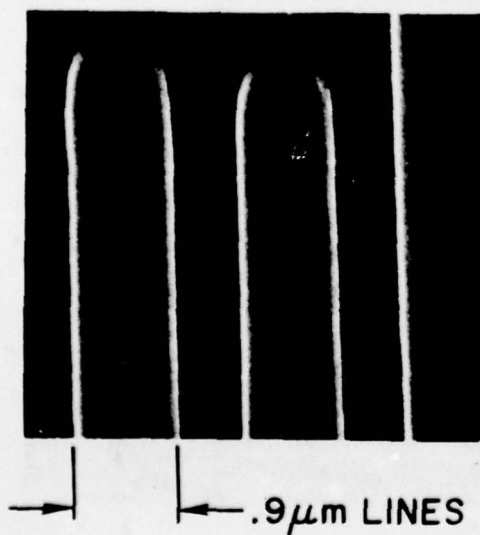
Fig.10 Photograph of submicrometer aluminum transducer lines made by direct optical projection
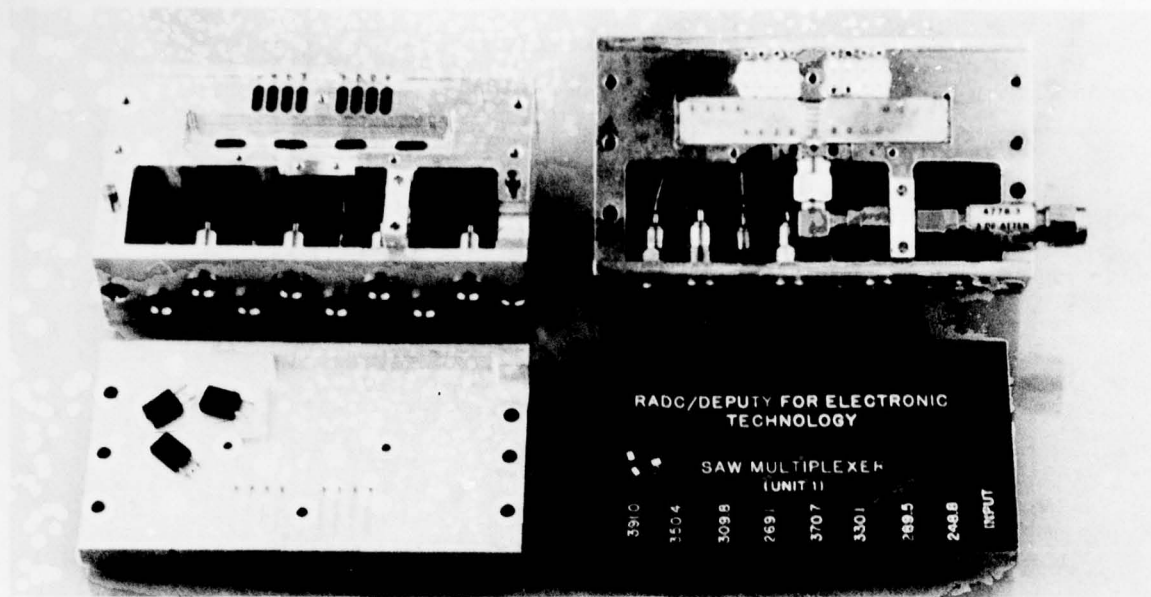
Fig.11 Photographs of a 1 x 8 UHF SAW multiplexer. The photograph on the left shows the empty can while on the right we see the SAW filters
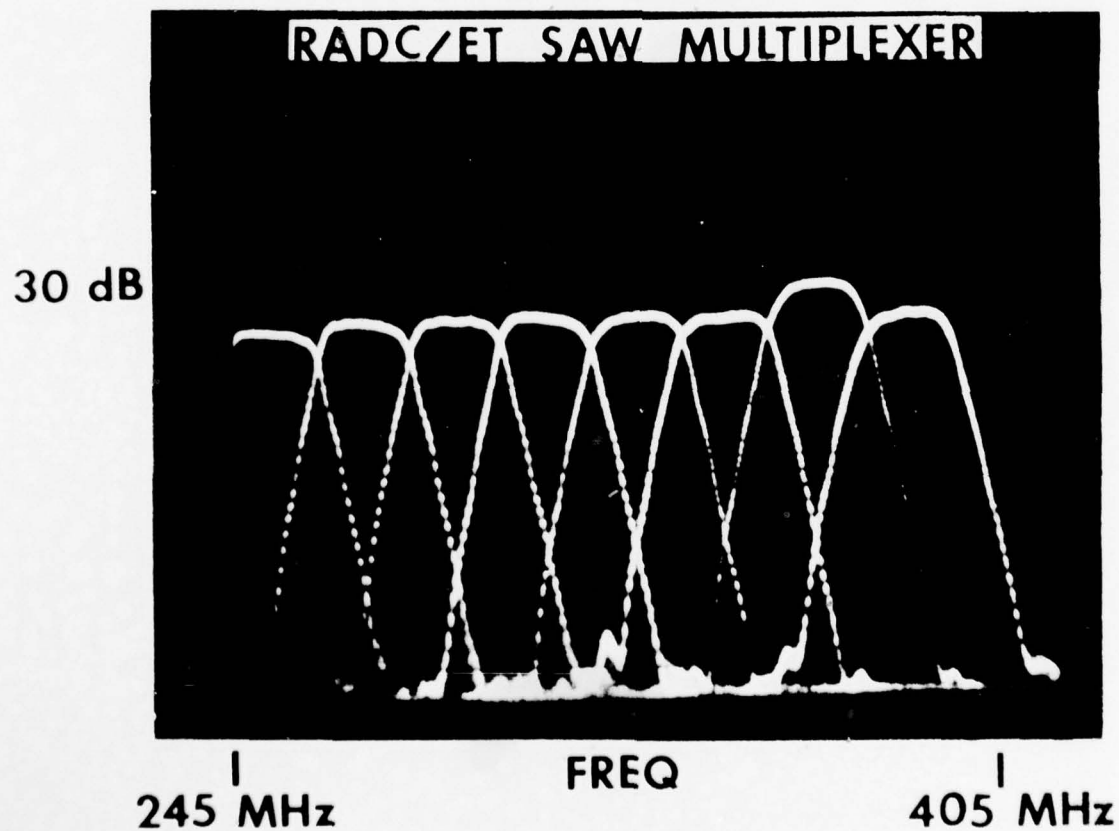
Fig.12 Plot of the insertion loss (10 dB/div) versus frequency (20 MHz/div) for the eight different outputs of the SAW multiplexer of Figure 11

30 MHz DELAY LINE OUTPUT

HORIZONTAL SCALE: 5 μ sec/cm

Fig.13   Comparison of the 9.1 μ sec delay line output for Y-cut, Z-propagating lithium niobate (left), and Minimum Diffraction Cut lithium tantalate, illustrating the lower volume waves of the latter

LOW SPURIOUS SAW DELAY LINE
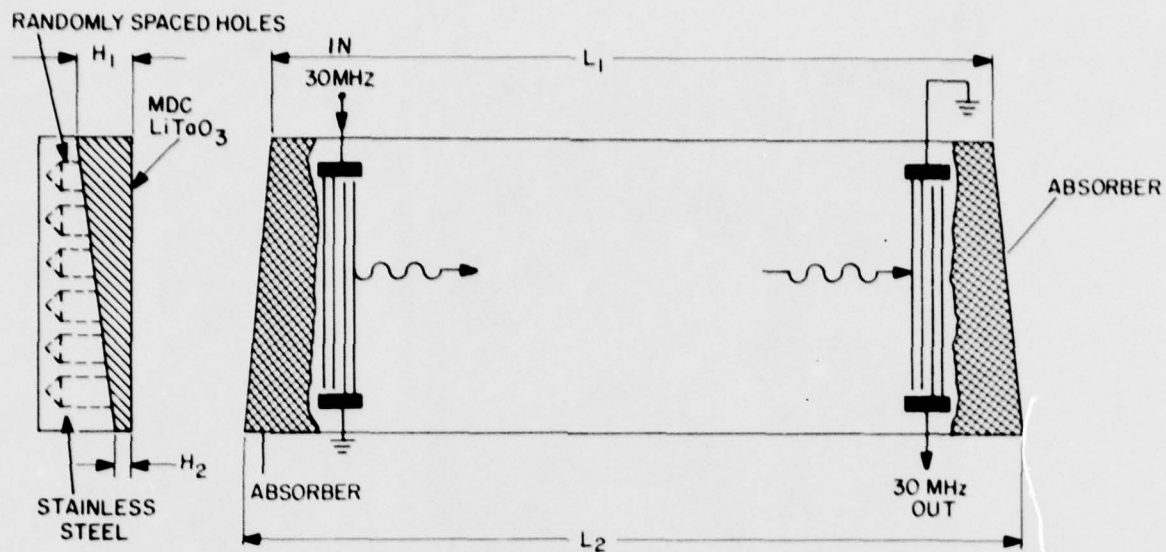
Fig.14   Line drawing of the 9.1 μ sec delay line showing the canted sides for suppression of bulk wave.

TABLE I

9.1 μsec DELAY LINE PERFORMANCE

	GOAL	THEORETICAL	EXPERIMENTAL
INSERTION LOSS (dB)	35	37.6	34.1 - 37.7
TRIPLE TRANSIT AND OTHER SPURIOUS (dB)	60 min	73.6	73.6
INPUT VSWR FOR 50 OHMS	2.0	2.0	2.1
DELAY V. S. FREQUENCY 29 - 31 MHz DEVIATION FROM LINEAR	$\pm 1.5^0$	$\pm .2$	$\pm 1.5^0$
TEMP. COEF. OF TIME DELAY (PPM)	100	64	MDC LiTaO$_3$
I dB BANDWIDTH (MHz)	3.0	2.4	2.3
CENTER FREQUENCY (MHz)	30.0	30.0	29.6
MAX. INPUT POWER (dBm)	20	20	20

THE MONOLITHIC INTEGRATION

OF SURFACE ACOUSTIC WAVE AND SEMICONDUCTOR CIRCUIT ELEMENTS

ON SILICON

FOR MATCHED FILTER DEVICE DEVELOPMENT*

Fred S. Hickernell
Motorola Government Electronics Division
Scottsdale, Arizona 85252

Henry J. Bush
Rome Air Development Center (DCID)
Griffis AFB, New York 13441

SUMMARY

This paper presents work on the development of silicon-based monolithic programmable matched filter SAW devices capable of generating and processing binary code data. Two programmable filters have been designed, fabricated and tested using zinc oxide film transducers for surface acoustic wave generation, piezoresistive MOSFET taps for detection, and semiconductor logic elements for phase and amplitude control of the taps. The filters were developed at a carrier frequency of 100 MHz, with a 10 MHz bandwidth and a capability for generating and correlating 31-bit PN sequences. The MOSFET taps on the first device were controlled by a 6 x 31 read-only-memory (ROM) fixed code library which could be randomly accessed. In the second device the ROM was replaced by a digital static shift register and a bucket brigade for arbitrary phase selection and amplitude control. The developments have served to demonstrate the viability of the monolithic integration of SAW and semiconductor elements, and permitted an assessment of various aspects of the technology.

1.   INTRODUCTION

Surface acoustic wave (SAW) components and devices have found wide application in electronic systems based upon their ability to perform basic analog signal processing functions in a more simple, reliable, and economical manner than their electronic counterparts. They are aptly suited for interfacing with electronic circuit assemblies: (1) as discrete packaged components with connectors or pin leads, (2) as chips on hybrid microelectronic circuits with wire and metallization interconnects, and (3) monolithically integrated with electronic components and devices on a common crystalline substrate. Each interface level has an important place in the application of surface acoustic wave technology to electronic systems.

It is the monolithic level of integration that provides the greatest opportunities and challenges. The future progress of SAW technology may well be measured by its ability to effectively compete with the myriad of innovative silicon semiconductor devices and logic modules, which will perform more complex signal processing functions at higher data rates. It is not believed that an improved silicon device technology will obviate the need for surface wave devices, but rather that a combination of the technologies can provide an even higher level of signal processing capability than one alone could do. For certain analog signal processing functions, there will be performance, size and economic advantages favoring the monolithic integration of surface acoustic wave and semiconductor device functions on a single silicon substrate. It is important to the future of both technologies to develop design capabilities, and improved material properties and fabrication processes which permit such integration.

This paper will present the results of two programs sponsored by Rome Air Development Center for the development of silicon-based programmable matched filter devices capable of binary code generation and detection. The matched filter devices developed under these programs have utilized zinc oxide film layer transducers for surface wave generation on the silicon and MOSFET devices for signal detection, and for phase and amplitude control. The developments have served to demonstrate the viability of such monolithic integration and permitted a quantitative assessment of particular aspects of the technology. This paper will describe the design, processing, fabrication and evaluation of the programmable devices and discuss advantages and limitations of the technology.

2.   GENERAL INFORMATION

2.1   Basic Device Approach and Performance Requirements

A pictorial representation of the three main functional parts of the basic device is shown in Figure 1. The various acoustic and electronic functions are developed on a properly oriented and doped silicon substrate using semiconductor process compatible-fabrication techniques. At each end of the device is a film layer transducer comprised of a thin aluminum interdigital electrode structure with an overlay of piezoelectrically active sputtered zinc oxide. The transducers are the electrical inputs for rf pulses in the encoding mode and the rf bi-phase sequences in the decoding mode.

Between the two transducers is a series of MOSFET device taps which serve as piezoresistive detectors of the surface acoustic wave energy generated by the transducers. These MOSFET cells are placed in geometric arrangements corresponding to the desired phase coding. Binary and quadraphase coding can be easily implemented. The device output is taken from the parallel connected drain lines of the MOSFET device taps.

The third major feature is the semiconductor circuitry controlling the phase and amplitude of the MOSFET detector taps. The information in this control circuitry may be statically or dynamically programmed. It codes and weights the taps according to the desired signal processing function.

The basic performance requirements for the two devices developed under the RADC programs were similar and are shown in Table 1.

Table 1. Basic Performance Requirements

Operating Frequency	100 MHz
Bandwidth (3 dB)	10 MHz
Modulation	Bi-Phase Pseudorandom
Number of Chips	31
Burst Length	3.1 μs
Chip Rate Within Burst	10 megachips/second

The first device was developed with phase programmability provided by an integrated read-only memory (ROM) which was mask programmable. The second device was developed for full phase and amplitude programmability using a digital static shift register (DSSR) and bucket brigade device (BBD) for the respective functions. The following section describes in more detail specific design and performance capabilities of the two programmable devices.

2.2   Device Design

In order to properly design the wavelength periodicities required by transducer and detector structures, the acoustic surface wave velocity characteristics of silicon with glass oxide layers must be determined. The Rayleigh velocity is dependent upon the elastic properties of the silicon and oxides, and upon the orientation of the silicon. On silicon the velocity is in the 4500 to 5100 m/s range. The thin layers of glass oxide required in fabricating and passivating the MOSFETs act to reduce by 3% to 5% the Rayleigh velocity on silicon. The velocity reduction is larger in the transducer structure due to the slower velocity of the zinc oxide. The velocity values for the transducer and MOSFET tap region can be obtained by computer calcuation but must be refined by experimental measurement.

A broadband low-loss means for conversion from electrical to acoustic modes is required for processing at a maximum information rate. It is desireable to have transducer structures which are compatible with semiconductor device processing techniques. The piezoelectric film layer transducer best meets these criteria. Sputtered zinc oxide film layer transducers have demonstrated capabilities for efficient excitation on silicon (HICKERNELL, F.S.., 1973). The transducer efficiency is geometry (KINO, G. S.., 1973) and film quality dependent (HICKERNELL, F. S.., 1976). Figure 2 shows the coupling factor dependency as a function of electrode geometry and film thickness-to-wavelength ratio. The k^2 coupling factor maxima vary from 1% to 3%, which permits a minimum loss device with bandwidths in excess of 10% to be developed. The transducer fabrication, which includes interdigital electrode delineation and vacuum sputtering of the zinc oxide, becomes an integral part of the overall wafer processing.

In the two devices developed, the transducers were designed for operation at 100 MHz with a 10 MHz bandwidth. ZnO film thicknesses of both 0.05λ and 0.4λ were used in maximum efficiency electrode geometries. On the first device development, quarter wavelength fingers and spaces were used with a thick film type I transducer and a thin film type IV transducer. On the second device a split electrode configuration was used in the thin film type II configuration.

The use of MOSFET piezoeresistive detectors (CLAIBORNE, L. T.., 1971) are very attractive from two perspectives. First, they have a large detection bandwidth, in excess of 100%. Second, detection phase and amplitude are easily controlled by gate voltage. The detection efficiency of MOSFETs has been examined theoretically and experimentally (STAPLES, E. J.., 1971 and DEFRANOULD, Ph.., 1972). The piezoresistively generated current on the drain line i_D is related to the power normalized gauge factor, Gp, the steady state dc drain current, I_D, the radian frequency of operation, ω, and the acoustic power per unit beamwidth P/W_B, by $i_D = Gp\ I_D\ \omega\ P/W_B$. Gp is a measure of the degree of variation of carrier surface mobility produced by the lattice stresses of a surface acoustic wave. It varies with crystal surface orientation, direction of propagation, and carrier type in the conducting MOSFET channel. Calculated values of the gauge factors for specific orientations of silicon are given in Table 2.

Table 2. Surface Wave Piezoresistive Gauge Factors for N and P Channel
Silicon MOSFET Structures (DEFRANOULD, Ph.., 1974)

Orientation		Gauge Factor 10^{-8} $(M \cdot S/W)^{1/2}$	
Silicon Cut	Propagation Direction	N-Channel	P-Channel
(001)	[100]	7.08	0.46
(001)	[110]	2.81	6.31
(110)	[110]	2.27	4.36
(110)	[100]	6.45	0.80
(111)	[110]	2.63	3.83
(111)	[112]	2.31	3.42

Analysis of the equivalent circuit of an N tap device for P_o/P_i (output power divided by input power), under conditions in which the total filter capacitance is parallel-resonated by an inductance, and where the internal load matches the output load and there are no parasitic resistive losses, yields for a single MOSFET tap,

$$\frac{P_o}{P_i} = \frac{I_D^2 \, Gp\omega^2}{4W_B g},$$

where g is the channel conductance. For a given gate-to-source voltage, it is desireable to operate the device in the saturation region to minimize g. The insertion loss per tap under typical operating conditions is in the 35 to 40 dB region.

The first device developed was fabricated on (110) n-type silicon with SAW propagation along the [110] direction and detection with PMOS FET cells. The second device was fabricated on (001) p-type silicon, with SAW propagation along the [100] direction and detection with NMOS FET cells. In each device, 31 paired MOSFET taps were developed, with one tap being offset from the other by half an acoustic wavelength.

The circuitry which controls the MOSFET tap phase amplitude can be external or integrated on the same chip. If control is external, a package or microstrip with 62 leads is required to address all taps. It is better to integrate the control circuitry on the same chip and minimize the external connections.

On the first of the two devices, a mask programmable read-only-memory (ROM) was developed containing six fixed 31-bit codes. The particular code was selected by biasing one of the six input lines, which in turn gate-biased the corresponding taps. The code was masked-programmed into the ROM by leaving either a thin gate oxide or a thick field oxide under the gate metal of each ROM driver. The ROM was designed to switch codes within 3.0 μs.

On the second device the control circuitry consisted of a 31-stage digital static shift register (DSSR) for tap phase selection, and a 31-stage bucket brigade device (BBD) for tap amplitude control. This represented a more flexible programming capability, with independent phase and amplitude control, at the expense of a more complex control circuitry situation. This type of device more closely approaches an ideal transversal filter. It represents the highest degree of integration of semiconductor and SAW elements attempted on a single silicon chip.

2.3   Mask Fabrication

The overall size of each device dictated that the mask layers required for the processing steps be composed. In each of the devices there was a major division into three cells, two end cells containing the transducers and the center cell containing MOSFET taps and control circuitry. A printout of one end of the ROM programmable device is shown in Figure 3. The repetitive nature of the taps and control circuitry permitted a further division into a basic cell that could be stepped and repeated. The masks were generated from computer tapes to an Electromask Stepper. The eight mask layers required for the shift register-bucket brigade controlled device were generated from a total of 48 individual 10x reticles. Four-inch emulsion masks were used with 40 and 60 devices respectively on each of the two device masks. Figure 4 shows the overall layout of a single DSSR-BBD device.

2.4   Wafer Processing and Characterization

Existing metal gate CMOS processes were modified to produce the required p-channel and n-channel circuitry for each device. Trade-offs between conductance, parasitic capacitance and resistance were analyzed and tested in order to specify the substrate doping and process parameters consistent with obtaining a high efficiency device. Thicker than normal field oxides were used to lower the drain current to substrate capacitance and insure proper coding operation in the ROM programmable device. A thick aluminum (15 KÅ) assured step coverage and

also reduced parasitic capacitance by permitting a narrower drain current summing line. The ROM programmable devices were fabricated on a 2-inch, 20-mil thick wafer with a capability for two columns of 18 devices each (see Figure 5). The DSSR-BBD programmable device was processed on a 3-inch, 20-mil thick wafer with three columns of devices and a capability for producing 38 device die per wafer (see Figures 6 and 7). The wafers were processed in lots of 10.

The zinc oxide transducer processing followed MOSFET wafer processing after glass passivation of the entire wafer surface. The ten finger pair aluminum interdigital electrode structure was patterned directly on the passivation glass. The zinc oxide was dc triode sputtered through a stainless steel mask where the interdigital patterns were developed. The aperture mask was offset from the wafer a few mils in order to develop a tapered transition region between film and glass passivation. Details of the sputtering conditions have been described elsewhere (HICKERNELL, F.S.., 1976).

Prior to die separation, wafers were visually and electrically tested to determine which devices were suitable for packaging. The bulk of the electrical testing was done on the Fairchild Sentry 600 Tester, which is capable of performing appropriate electrical measurements on the MOSFET circuitry of all dice on the wafer and of printing out the results. A specially made probe card was used to contact all input and output pads of the circuit. Among the various parameters evaluated, the tests identified devices with high leakage currents, determined tap current uniformity and tested functionality of the control circuitry (e.g., bucket brigade and shift register devices). There was an excellent correlation between the results of the wafer testing and the rf performance. The yield of good devices averaged near 18% for the ROM devices and exceeded 40% for the DSSR-BBD devices.

The final selection of dice was made on the basis of visual examination of the electrically good dice. A Tempress saw was used to separate the long narrow shaped chips. The die sizes were 770 mils by 75 mils for the ROM controlled device and 860 mils by 160 mils for the DSSR-BBD controlled device.

2.5   Device Assembly and Packaging

A commercially available 1.0 inch by 0.5 inch all-metal flat-pack from Isotronics Inc. was chosen to house the MOSFET chip and ceramic thick film interconnect substrates. The chip was placed off-center to keep the wire bonds to the drain line and ZnO transducers as short as possible. All internal package interconnections are 1 mil aluminum wire ultrasonically bonded. The gold metallized ceramic substrate and the MOSFET chip were secured in the package using a conductive silver adhesive (E-KOTE 3030). The adhesive provides a strong bond with high electrical conductivity while permitting part removal by dissolving the adhesive in Trichloroethylene. Dow Corning RTV 3144 was used to dampen the unwanted end reflections. A brass test fixture with OSM connectors was used to house the packaged device and matching circuit during rf testing. Special care had to be taken to electrically isolate the matching circuitry so as to prevent any direct feedthrough of signals. Photographs of the packaged ROM and DSSR-BBD devices are shown in Figures 8 and 9.

2.6   RF Evaluation

Testing at rf can be done at two levels. The first is at the basic device level, which characterizes performance parameters and determines if the device is suitable for testing as a system element. Testing at a system level involves functioning of the device under conditions in which it will be used, such as continuous signal encoding and decoding in a simulated communications environment. Most of the testing of the devices has been done at the basic device level. A short (100 ns) burst of pulsed rf applied to the packaged and matched devices is used to determine the efficiency of transducer and taps, characterize the phase and amplitude of the output waveform, and determine the signal feed-through level. The transducer-to-transducer loss corrected for attenuation through the taps gives a measure of the transducer efficiency. The loss from the transducer to the first and subsequent taps can be used to establish the MOSFET detection efficiency and propagation loss through the MOSFETS. The short pulse also isolates the direct feedthrough signal from the detector output signal, and its magnitude may be measured.

Figure 10 shows the output from 31 in-phase PMOS FET taps in response to a pulse input. The total loss from pulsed rf to output of the first tap is approximately 80 dB under a nominal drain current of 15 mA. The single transducer loss is typically 15 dB with an added 3 dB for bidirectionality and 3 dB for acoustic channel splitting. The expansion loss for 31 bits is 15 dB and there is typically an additional 3 dB loss for matching and parasitic elements. This gives a piezoresistive coupling efficiency near 40 dB per tap. For the NMOS FET taps a coupling efficiency of 35 dB was obtained under similar operating conditions.

It is evident from the scope photograph that there is a substantial signal amplitude droop and a coupling pulse which exceeds in amplitude the output signal. These two factors will be addressed in the section which follows on technology analysis.

If the device taps are encoded by ROM, DSSR-BBD or hard wiring, correlation properties can be evaluated. The code generation can be from the matched device or a separate electronic code generator. Figure 11 shows a series of three encoded waveforms on the ROM programmable device. Because amplitude droop limited the coded outputs of the devices developed, a better measure of correlation properties was obtained by generating a uniform amplitude code electronically. The best performance was obtained with a thin ZnO transducer (2.8 μm) on the ROM controlled device, which had only a 6 dB coded waveform amplitude droop. A photograph of the

autocorrelation output waveform for a 31-bit code is shown in Figure 12. The peak-to-sidelobe autocorrelation was 14.5 dB compared to a theoretical of 17 dB. The correlated peak was very symmetrical and had the theoretical 200 ns width. The auto to cross-correlation margin was 9 dB. The loss from the expanded sequence to the correlated waveform was 44 dB, consistent with the 40 dB/tap piezoresistive conversion loss. The photograph indicates a high signal quality for the correlated output.

The utility of the ROM device for systems use was demonstrated by generating long coded sequences in a pseudorandom manner. Two of the ROM controlled devices with supporting logic, rf switching and amplifier circuits were randomly coded and alternately pulsed to generate the continuously coded signal. The logic used could generate a short 64 word nonrepeating or a 5-day non-repeating sequence.

The DSSR-BBD device was not evaluated in a system environment. The ability of the digital shift register and bucket brigade to control tap phase and amplitude was evaluated. The shift register would control the tap phase but with some degradation in amplitude when a single tap was phase switched. This was due to off-on ratios in the shift register circuitry. The range of tap amplitude control through the bucket brigade was 10 dB. The desired amplitude control of 20 dB was limited by body effect, the change in threshold voltage with bias which occurs in NMOS FETs. The ability to control tap phase and amplitude was demonstrated and design changes can correct present deficiencies to obtain a full range of tap control.

2.7   Technology Analysis

A technological analysis of this type of SAW programmable device may be looked at from both fabricational and performance viewpoints. Fabricationally, the device is large and unusual in size but standard LSI circuit technology can be directly applied. There is not as high a density of active devices as in some LSI circuits. The packaging and assembly techniques are considered standard. Correlation between wafer probe and rf device evaluation has been excellent. Yield factors are high, in excess of 40% on the second device. In quantity production with proven mask sets, an even higher yield is anticipated. All evidence points to the fact that in large quantity production a very low device cost can be attained.

The overall performance of the ZnO/Si-MOSFET matched filter devices is considered extremely encouraging. The MOSFET tap generated and detected waveforms on both devices were clean and phase reversals were distinct. Autocorrelated waveforms have approached theoretical values. A high degree of rf tap-to-tap isolation has been achieved.

Two major problem areas were encountered on the ROM programmable device: 1) direct signal coupling between input and output, and 2) considerable amplitude droop in the time wave-form. The signal feedthrough was only 70 dB below the input pulse under the best grounding and isolation conditions that could be obtained. In the DSSR-BBD programmable device, additional design measures were taken to assure that there was no feedthrough paths internal to the silicon or its immediate package environment. The direct signal coupling was brought to the 90 dB level and it is not considered an inherent device problem.

The amplitude droop problem which was also present on the DSSR-BBD device, is related to the standard MOSFET structure. Step oxide and aluminum regions are present with structural discontinuities up to 2 μm in the path of the surface wave. Even though these represent small fractions of a wavelength, (<0.05) the additive effect of several repeated discontinuities is to cause a significant amount of surface wave energy to be reflected. The solution to this problem is to minimize these structural discontinuities. The height-to-wavelength ratio of the discontinuities must be reduced to near 0.01 to bring the total reflection through 31 taps to the 1 dB level. This is not easily achieved through modification of standard processes. Decreased oxide and aluminum thicknesses must be traded off against increased parasitic capacitance and resistance, and current carrying ability. It is estimated that signal droop across a 31 tap structure will be brought to 3 dB by adjustments in standard process techniques. Additional improvement must come from isoplanar processing with buried oxide layers and silicon contacts to the source and drain regions. Analysis and experimentation will be required to confirm these expectations.

The addition of the bucket brigade and shift register as control circuitry for the taps increased the complexity of the device and its sensitivity to process variables. The shift register presented no major problems and increased the versatility of code selection. The bucket brigade lacked sufficient dynamic range to have full tap amplitude control. A 20 dB dynamic range is possible; however, it may be necessary to consider other alternatives for tap amplitude control.

System-wise, the rate of coded data flow is limited by the time in which the taps may be reset by the control circuitry. In the case of the ROM controlled device, 3.0 μs was sufficient and it was possible to use two devices for contiguous coded signal generation. The DSSR-BBD control circuitry requires approximately 35 μs for the setting of the 31 tap phases and amplitudes. The bucket brigade is designed to operate at a 1 MHz rate and the shift register at a 10 MHz rate. By developing a 20 MHz shift register rate and using parallel inputs for voltage control, the coded data rate can be brought to near 3.0 μs.

Increasing the device efficiency should be a major area of concern in future work. There is room for improved zinc oxide transduction efficiency in the thick film mode. The MOSFET conversion efficiency could be improved through the use of a higher concentration of acoustic power in the channel region. This would help reduce the amount of drain current required to obtain useable signal levels. Presently the power per coded bit at maximum efficiency is in the 50 to 100 milliwatt region.

The use of piezoelectric gate oxides may be a viable approach for significantly reducing the conversion loss at the piezoresistive taps (GREENEICH, E. W.., 1975).

3.   CONCLUSION

The devices developed under the two RADC programs served to demonstrate the feasibility of melding piezoelectric film transducer technology and semiconductor MOSFET technology to achieve a monolithic programmable surface-wave matched filter with stored coding functions. There are several unique features which set these developments apart from other programmable matched filter developments. It is truly monolithic, using a silicon substrate upon which the MOSFET detecting structures and control circuitry are semiconductor processed and the zinc oxide transducer structures are developed. Film metalization interconnects are used throughout. The use of gate voltage controlled MOSFET taps for code phasing and amplitude weighting eliminates the necessity for complex hybrid switching structures which require considerable wire bond interconnections. A high degree of rf tap-to-tap isolation is achieved. The zinc oxide film transducers permit a wideband signal to be generated with relatively low loss. The bandwidth characteristic of the MOSFET detectors is extremely broad, and detection efficiency increases with increasing frequency. The reliability and reproducibility of the device, which uses standard MOSFET processing, is good and the potential cost in high volume should be substantially better than hybrid or more fabricationally complex epitaxial structures. The temperature coefficient of delay should be reasonably good because of the relatively low negative coefficients of zinc oxide and silicon, and the positive coefficient of SiO_2. Compared to digital electronic technology, it represents a simpler device structure. It presently requires somewhat higher power per bit than electronic counterparts but offers the advantages of operating directly at rf frequencies.

Work remains to be done in the minimization of tap reflections, increasing overall device efficiency and optimizing the tap control circuitry before this type of device reaches its full potential and is ready for system implementation. Present problem areas which restrict performance can be solved by design and process changes. Considering the amount of signal processing that can be done in such a small form factor at low cost, it seems important to develop new designs and better processing techniques, which capitalize on the best features of the technology and optimize the acoustic/semiconductor interactive effects.

4.   ACKNOWLEDGEMENTS

5.   REFERENCES

CLAIBORNE, L. T., STAPLES, E. J., and HARRIS, J. L., 1971, "MOSFET Surface Wave Detectors for Programmable Matched Filters", Appl. Phys. Letters, vol. 19, pp. 58-60.

DEFRANOULD, Ph., 1972, "100 MHz MOSFET Detection of Rayleigh Surface Waves", Proceedings 1972 IEEE Ultrasonics Symposium, pp. 229-232.

DEFRANOULD, Ph., 1974, "Transducteur d'Onde Acoustique de Surface a Effet de Champ", Thomson CSF Internal Report.

GREENEICH, E. W. and MULLER, R. S., 1975, "Theoretical Transducer Properties of Piezoelectric Insulator FET Transducers", J. Appl, Phys., vol. 46, pp. 4631-4640.

HICKERNELL, F. S., 1973, "DC Triode Sputtered Zinc Oxide Surface Elastic Wave Transducers", J. Appl. Phys., vol. 44, pp. 1061-1071.

HICKERNELL, F. S., 1976, "Zinc-Oxide Thin-Film Surface-Wave Transducers", Proc. of the IEEE, vol. 64, pp. 631-635.

KINO, G. S. and WAGERS, R. S., 1973 "Theory of Interdigital Couplers on Nonpiezoelectric Substrates", J. Appl. Phys,, vol. 44, pp. 1480-1488.

STAPLES, E. J., 1971, "Detection of Rayleigh Surface Waves on Silicon with the MOSFET", PhD. Dissertation, Southern Methodist University.
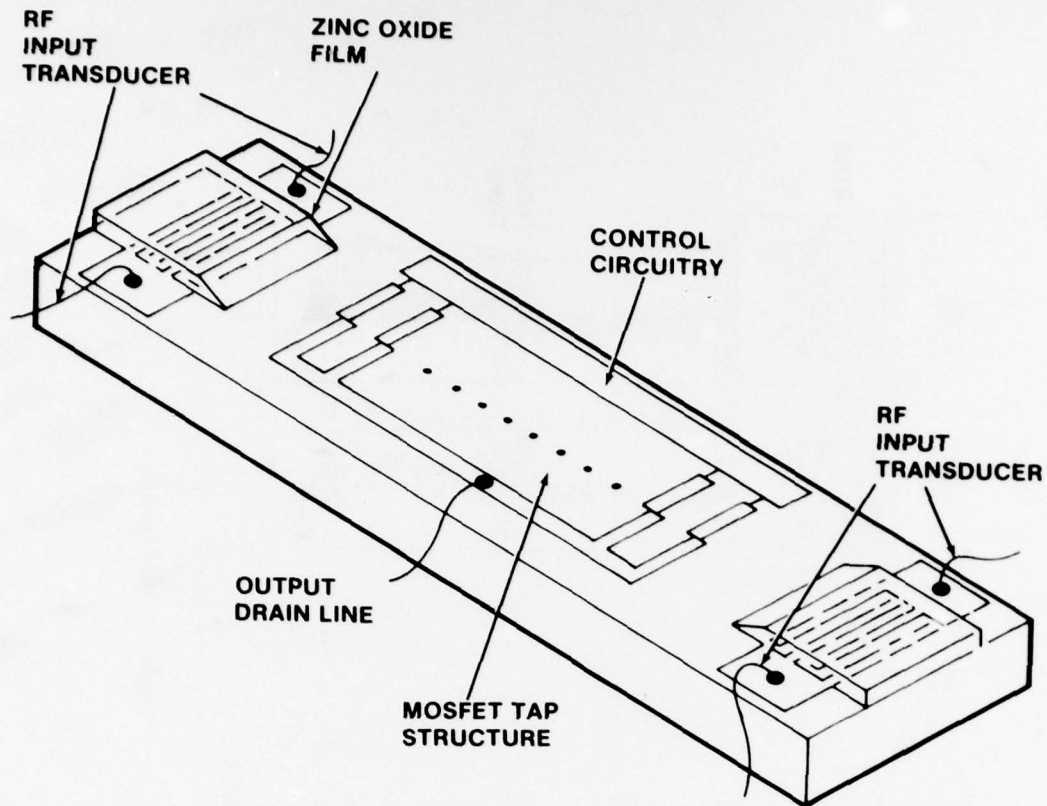
Figure 1. Monolithic ZnO/Si-MOSFET Programmable Surface Acoustic Wave Matched Filter
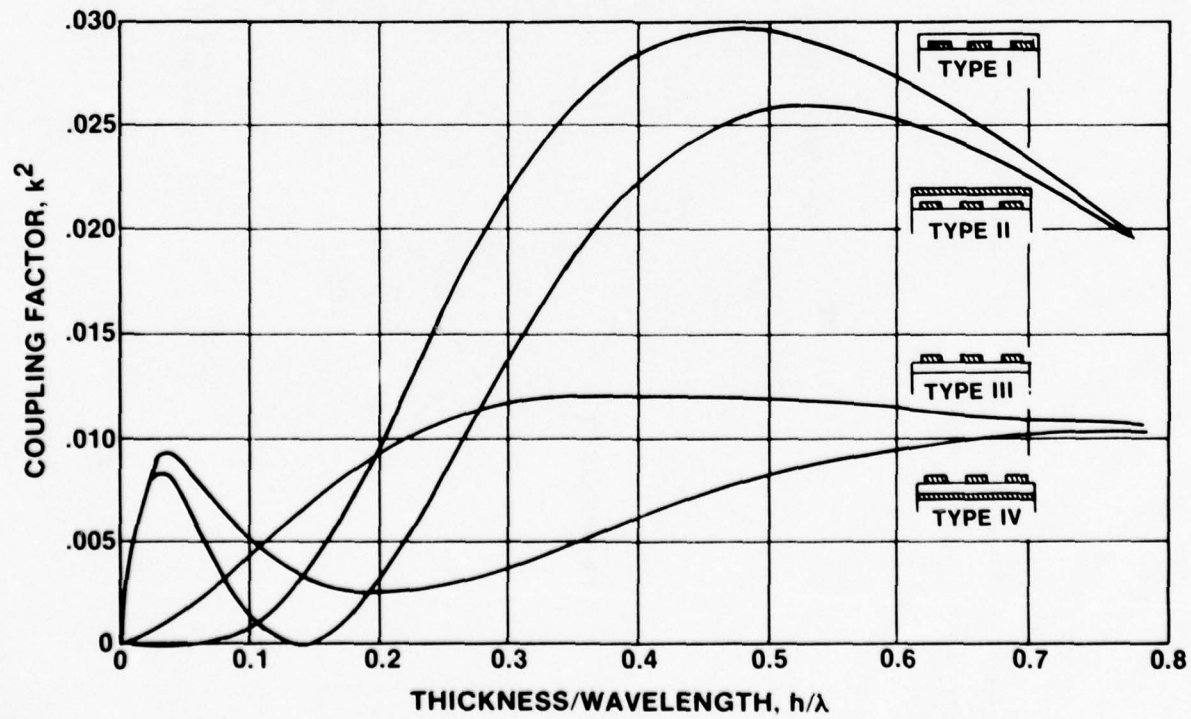
Figure 2. Coupling Efficiency of ZnO/SiO$_2$/Si Film Layer Transducer
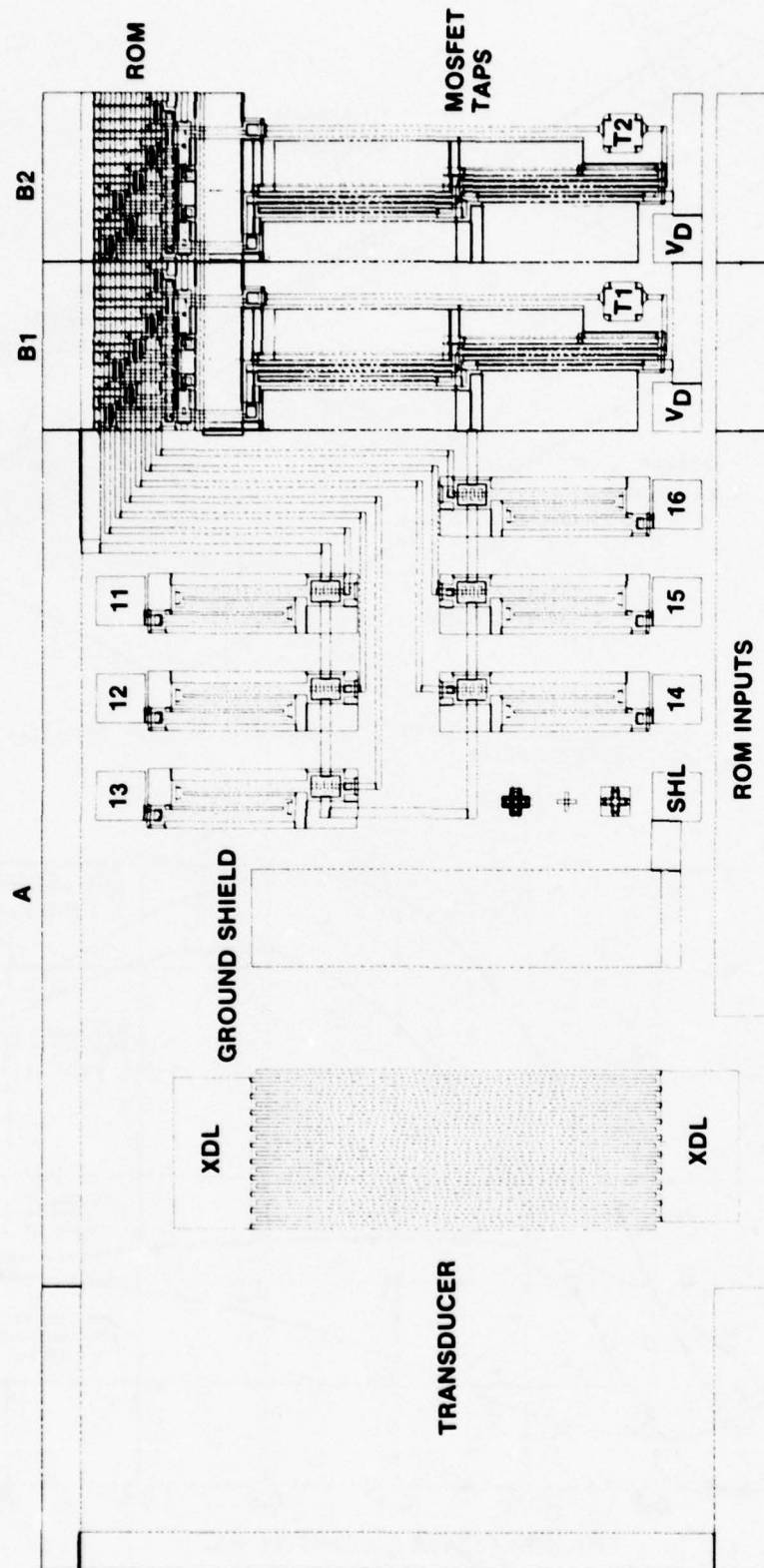
Figure 3. Layout of Programmable Filter at Transducer End Showing ROM Inputs and Two Tap Cells

DIGITAL STATIC SHIFT REGISTER

SAW TRANSDUCER

BUCKET BRIGADE DEVICE

MOSFET TAPS

SAW TRANSDUCER
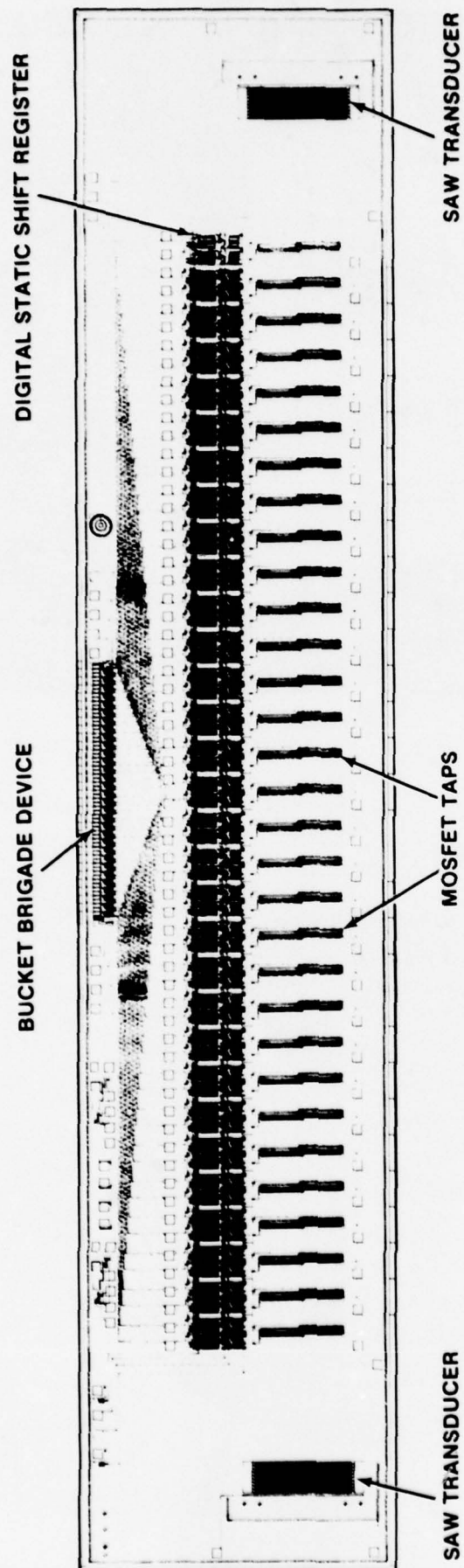
Figure 4. Layout of DSSR-BBD Programmable Filter
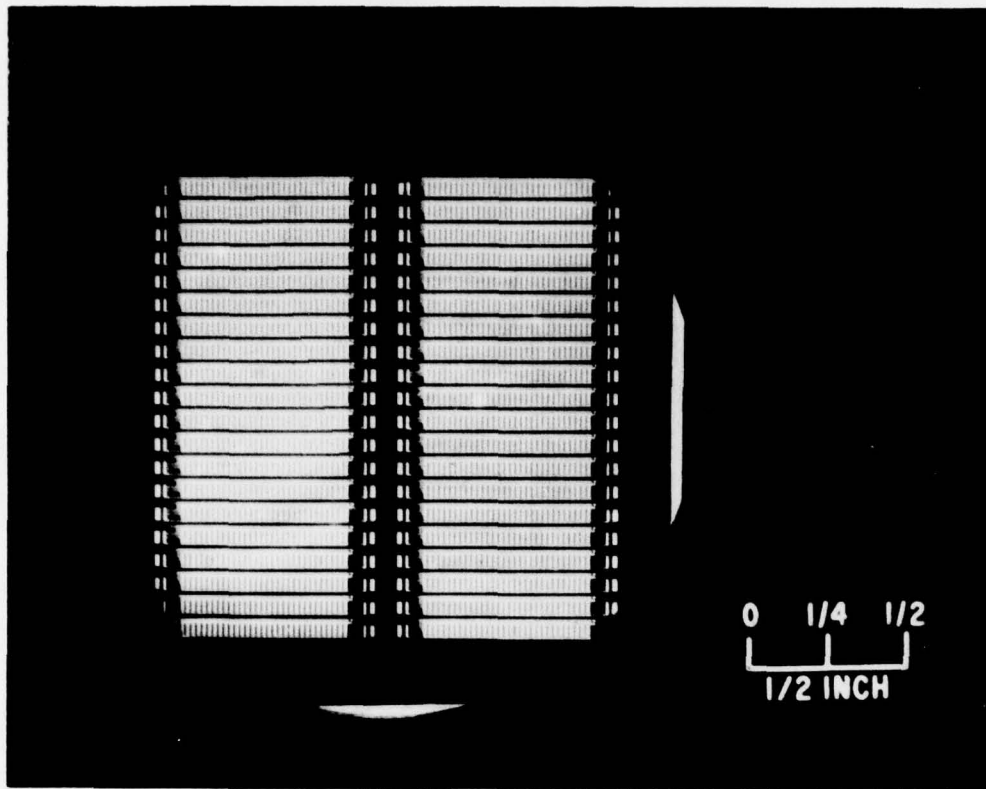
Figure 5. ROM Programmable Filter Device Wafer
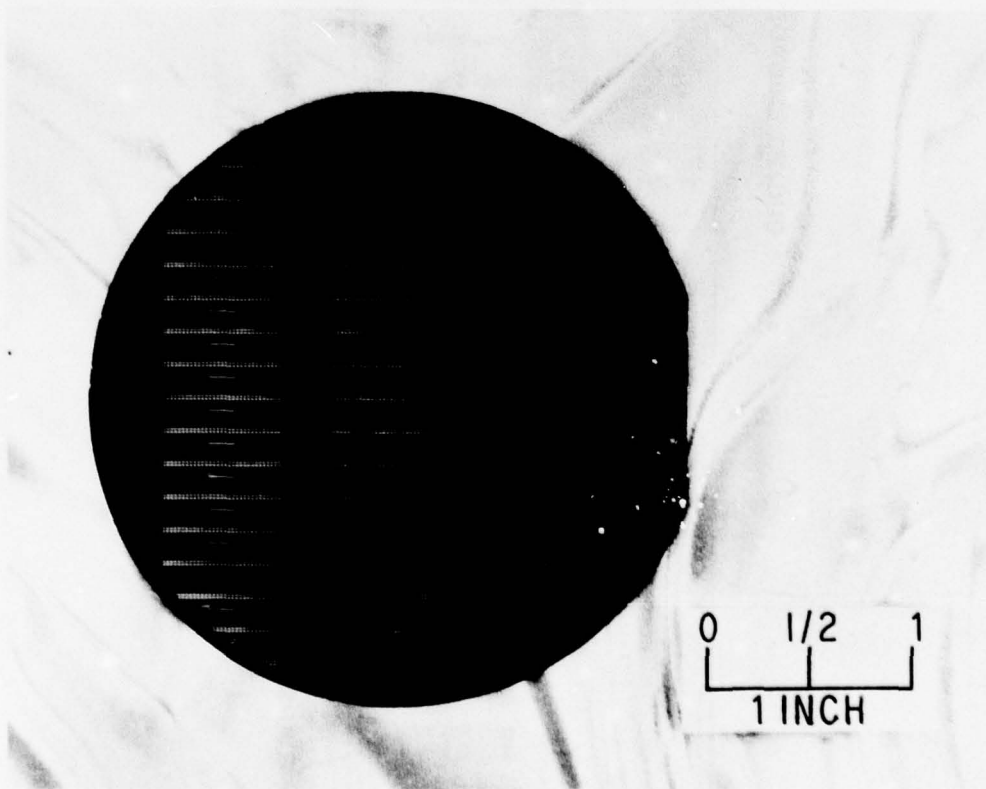
Figure 6. DSSR-BBD Programmable Filter Device Wafer
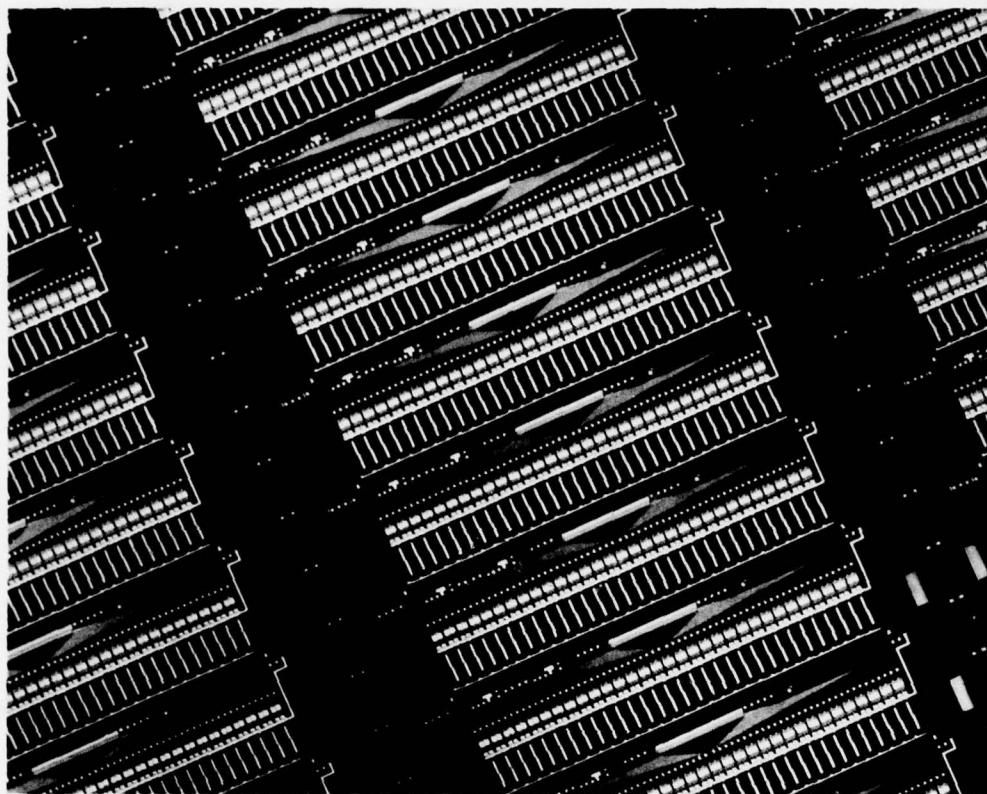
Figure 7.   Closeup of DSSR-BBD Programmable Filter Wafer
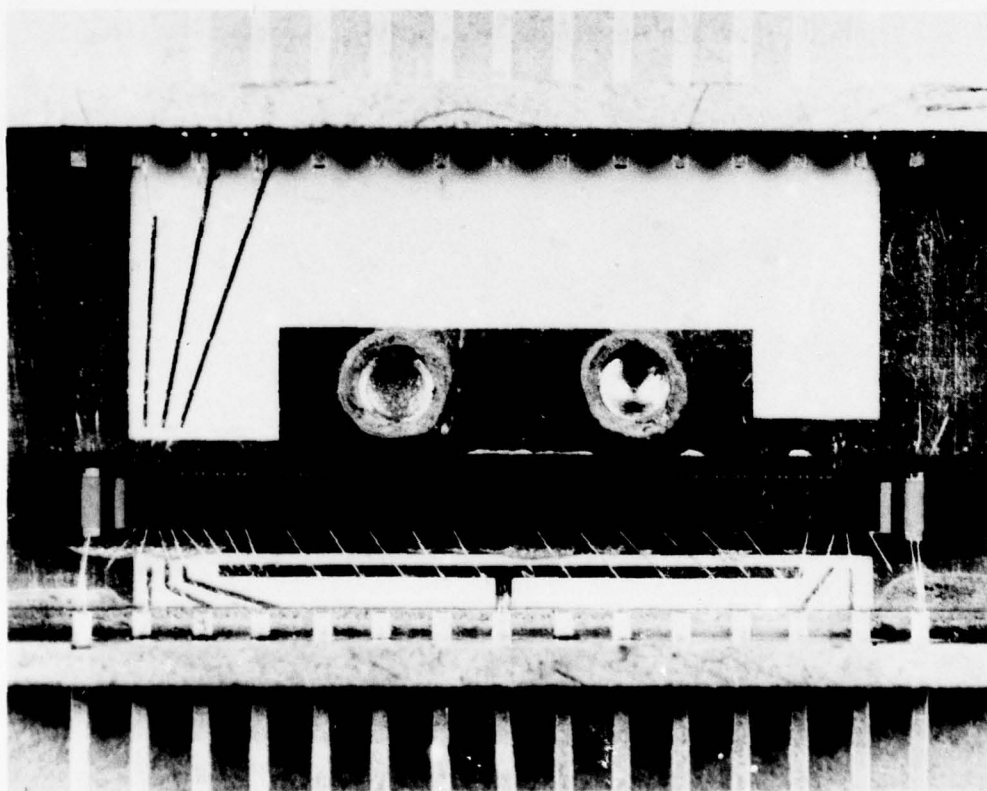
Figure 8.   Packaged ROM Programmable Transversal Filter
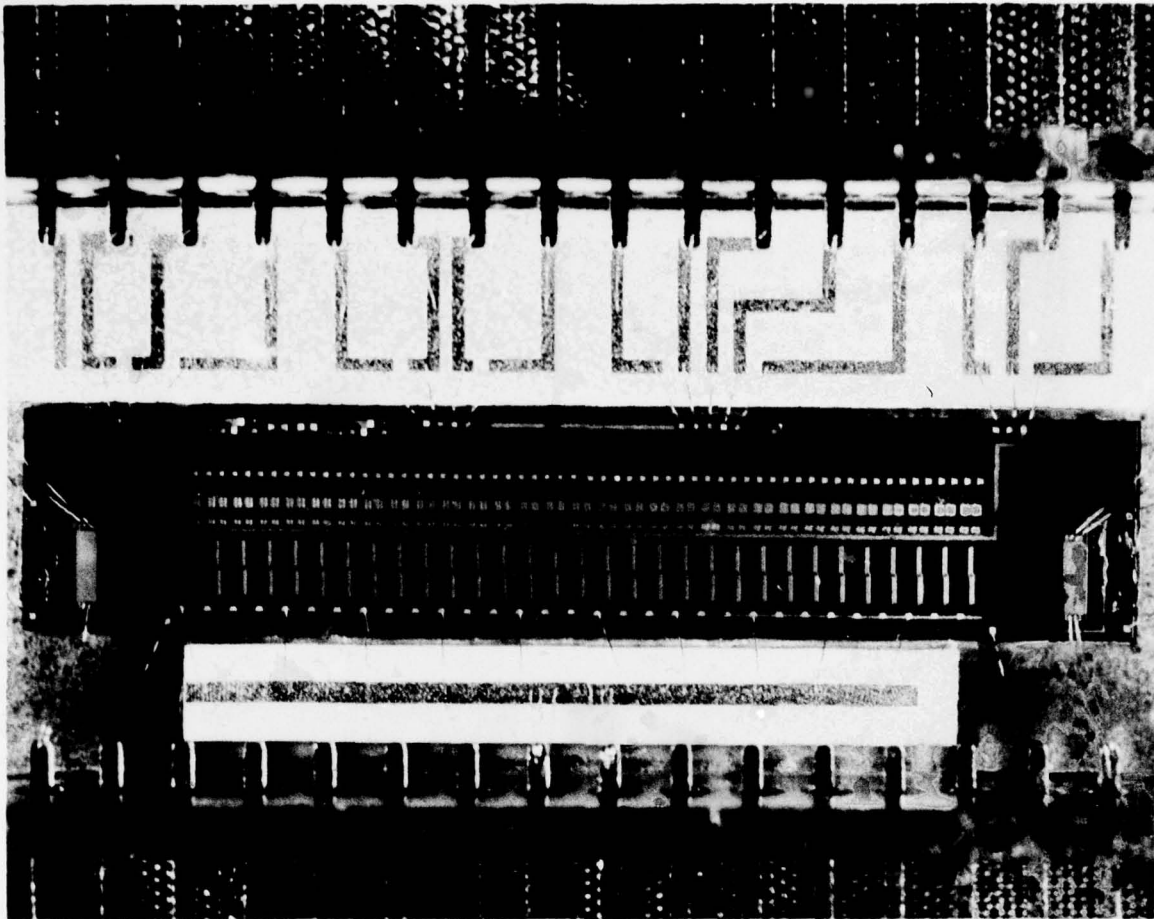
Figure 9.   Packaged DSSR-BBD Programmable Transversal Filter
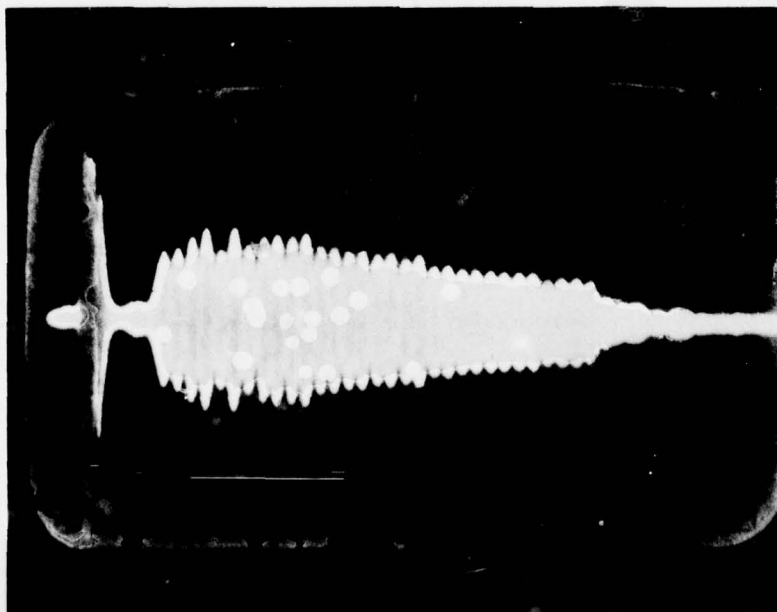
Figure 10.   Output from 31 Channel MOSFET Taps
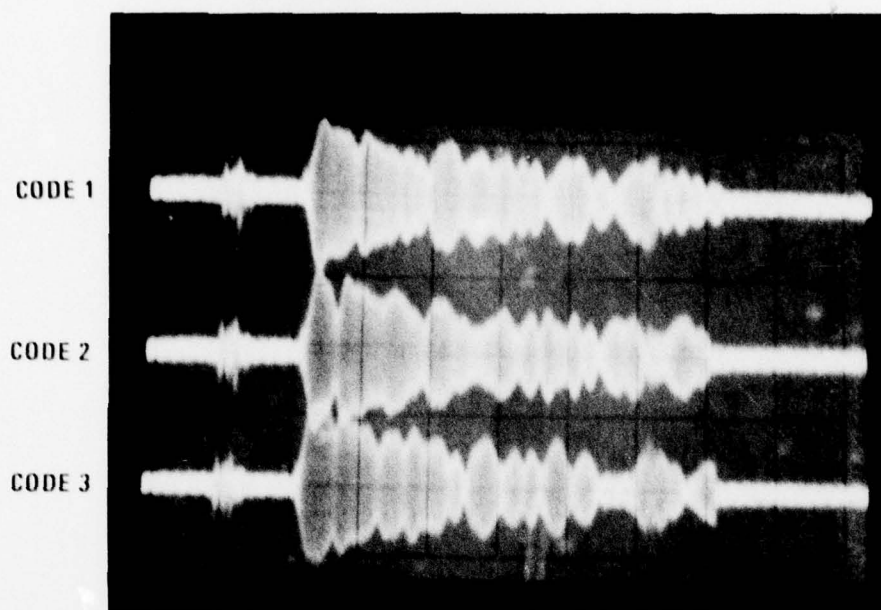
SCALE
HOR: 0.5 μs/DIV
VERT: 0.1 V/DIV

Figure 11.   Encoded Waveforms of the ROM Programmable Filter
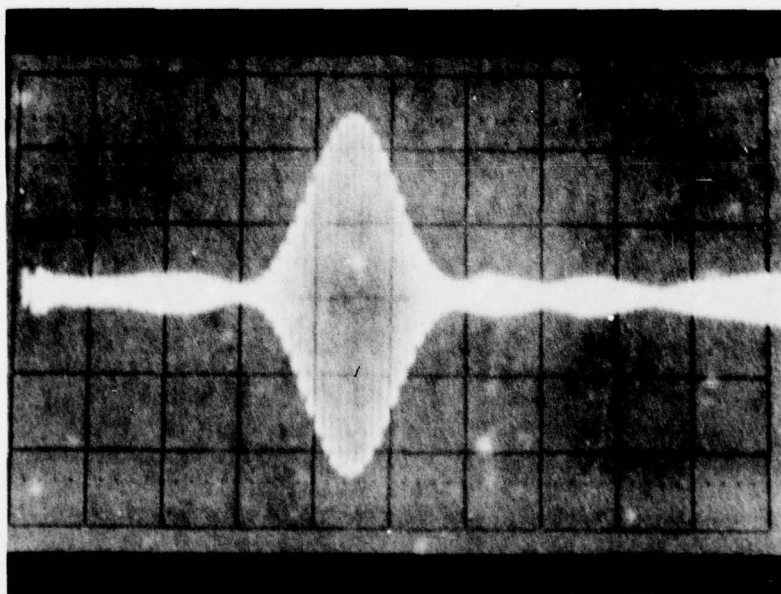
Figure 12.   Correlated Waveform for the ROM Controlled Programmable
Filter with a Digitally Generated Code Input

EXPERIMENTS AND ANALYSIS OF ACOUSTOELECTRIC MEMORY CORRELATORS

K.A. Ingebrigtsen, A. Rønnekleiv and S. Stueflotten
Electrical Engineering Department
The University of Trondheim
7034 Trondheim-NTH, Norway

SUMMARY

This paper describes progress in the development of acoustoelectric diode memory correlators. The fabrication technique is described in some detail and experimental results from devices using Schottky diodes and pn-diodes are given. A simple model for analysis of the performance is given. This model describes adequately several of the experimentally observed features in the recording process, of the storage, and of the correlation read out. However, it appears inadequate to explain an experimentally observed frequency dependence of the correlation output efficiency. The experiments suggest that the model needs improvements in the description of the recording process.

1. INTRODUCTION

The acoustoelectric diode memory correlator first demonstrated by Ingebrigtsen et al. (1975) emerged as a practical implementation of the concept of a fast programmable analog signal processor proposed by Stern and Williamson (1974). The basic concept is shown schematically in Fig. 1. It consists of a surface wave delay line with a silicon strip mounted adjacent to the delay line. The silicon surface facing the piezoelectric delay line crystal is covered with an array of free standing diodes. A signal launched into the delay line can be recorded and stored as charge on the diode array by applying a short impulse to the plate contact to forward bias the diodes. The diodes will charge up proportionally with the local field which is the sum of the space uniform field from the plate impulse and the local field from the surface wave. Thus, provided the plate impulse is short compared to the period of the rf-signal, the diode charges reflect the local phase and amplitude of the surface wave signal. The storage of the signal is determined by the leakage current through the diodes.

Through the nonlinear capacitance of the diodes, the spatially varying charge will mix with other signals of the same spatial period subsequently launched into the delay line and generate an output signal across the plate contacts which is essentially proportional with the correlation of the input signal with the stored charge pattern. Specifically, when the stored charge pattern is

$$Q(z) \cdot \cos kz , \tag{1}$$

and the external signal is

$$D(z-vt) \cdot \cos(\omega t - kz) , \tag{2}$$

the output voltage is

$$V_{out}(t) \sim \cos \omega t \int_{\substack{\text{interaction} \\ \text{length}}} Q(z) \cdot D(z-vt) dz \tag{3}$$

The initial experiments applied platinumsilicide Schottky diodes on 10-50 Ωcm n-type silicon (INGEBRIGTSEN 1976). These experiments showed that the diodes could be charged with a few nanosecond plate pulse and that storage times of tens of milliseconds could be obtained. Subsequent experiments using pn-diodes (DEFRANOULT et al. 1976) have shown different results, namely charging times of $\sim$ 0.1-1.0 µsec and storage times of 10-100 seconds. With these long charging times it is necessary to apply an rf-burst at the signal frequency to the plate to maintain coherence during the charging period of 10-100 rf-periods. Modified configurations and other permutations of input and output ports have been demonstrated (BORDEN et al. 1976, TAKADA et al. 1976, DEFRANOULT et al. 1976). In general the results show that the signal to be recorded must be a small signal superimposed on the large recording signal if a fast and efficient recording without distortion is desired. An interesting mode of operation is the integrating correlator (RALSTON et al. 1977, BORDEN and KINO 1977). The input signal is then applied both to the input transducer and the plate terminal. A delay line is inserted into the plate terminal to compensate for the delay from the input transducer to the center of the interaction region. If both signals are small signals they produce a stationary charge pattern on the diodes with a spatial period corresponding to the acoustic wavelength. This charge pattern is proportional with the autocorrelation of the input signal. By appropriate adjustment of parameters such as pre-biasing the diodes, the accumulated diode charge can be integrated for milliseconds. The stored charge pattern can finally be read out by applying a short scan pulse to the delay line. It is worth noting that although the autocorrelation gain corresponds to integration for milliseconds it is only the central part of the autocorrelation corresponding to the delay time through the interaction region that is recorded.

Since the first demonstration, notable progress in the fabrication technology has been reported. Although the first devices showed quite promising performance, they exhibited several undesired features such as uncontrollable electromagnetic feed-through, uncontrollable and often high spurious signal levels, and non-uniform interaction and storage across the length of the silicon strip. Substantial improvements have been reported, but further

improvements are probably desirable. An example of a reasonably well performing Schottky-diode device is shown in Fig. 2. The photograph shows the autocorrelation of a linear FM-chirp with a time bandwidth product of 150. The correlation is performed after a storage time of 3 milliseconds. The response is relatively clean with all spurious and feedthrough more than 25 dB down over the 30 MHz bandwidth of the chirp.

2. DESIGN CONSIDERATIONS AND FABRICATION TECHNIQUES

The fabrication of air gap coupled acoustoelectric devices can be divided in 4 main steps, silicon fabrication, delay line fabrication, assembly, and package design. The memory correlator imposes specific requirements on fabrication which differs from those of the acoustoelectric convolver (REIBLE et al. 1976). Most of our effort has been devoted to the delay line fabrication and the assembly technique.

The bulk resistivity of the silicon is usually chosen as a compromise between efficient interaction over the processing length, uniformity in the interaction, and uniformity in storage. As with the convolver, increased interaction leads to stronger acoustoelectric losses. In convolvers the optimum coupling therefore occurs for an acoustoelectric attenuation of 8.5 dB over the interaction length. This can be obtained for various silicon resistivities by adjusting the air gap correspondingly. However, for given center frequency, bandwidth, and interaction length, there is one specific combination of silicon resistivity and air gap which gives an optimum coupling and a minimum frequency variation over the bandwidth. An attenuation of 8.5 dB over the interaction length is often unacceptable in a correlator since it effectively leads to a reduction in processing gain.

Adding all these constraints one usually ends up with silicon resistivities between 10-30 Ωcm for devices with center frequencies in the range 100-300 MHz, and integration time below 10 μsec. Within this resistivity range it is our experience that the low resistivity samples generally show better uniformity over the interaction length. On the other hand, we have observed the longest storage times on high resistivity devices. Typically figures which can be related to the silicon are interaction uniformities better than 1-2 dB over 5-10 μsec interaction length, storage time variations by a factor of 2-3 over the inter-action length, and average storage time variations between samples from different wafers and different resistivities of a factor of 5. No doubt these figures can be improved by a more dedicated silicon technology than ours.

The delay line fabrication and assembly technique that we have explored are with some modifications based on a technology developed by the Microsound group at the MIT Lincoln Laboratory (SMITH 1975, and RALSTON 1975). One of the constraints encountered in memory correlator design which is not present in convolver design is to reduce the thickness of the LiNbO$_3$ crystal so that the diodes can be charged from a reasonable plate pulse source. Although the charging of the diodes primarily is controlled by the current, it is neces-sary that a voltage of 0.5 V or so is available to the Schottky diodes if they shall be charged with a single pulse. Since the applied plate pulse appears across the diodes in series with the capacity of the air gap and the LiNbO$_3$ crystal it is clearly desirable to make the LiNbO$_3$ crystal thin to reduce the requirement for available voltage from the plate pulse source. A schematic of the configuration is shown in Fig. 3.

To achieve this we have worked with LiNbO$_3$ crystals of thicknesses between 100-200 μm. The thin crystal is glued onto a thicker 500 μm LiNbO$_3$ crystal which also supports the backing electrode on the interface between the two crystals. For the bonding of the two crystals we have used a UV-hardening epoxy resin. The epoxy is spun on the base crystal in a conventional photoresist spinner to obtain a uniform layer of $\sim$ 3μm thickness. The two crystals are pressed together in vacuum to provide a uniform contact with no air bubbles in the epoxy layer. The epoxy is ultimately hardened by exposing it to UV-light. The transmission characteristic of a 12 μsec delay line on this configuration shown in Fig. 4, demonstrates a clean response without any observable multimode effects. The observed ripple is caused by reflections from an auxiliary transducer.

In the convolver, where the output is on the second harmonic of the input signals, all spurious modes can be effectively filtered at the output. This is not the case with memory correlators where the output occurs on the signal frequency. Thus all spurious signals appear in the passband of the output signal. Since the correlation output usually is 40-60 dB below the input signal, further suppression of spurious signals puts severe demands on the design. With the described configuration we observe spurious signals which can be related to thickness resonances of the thin LiNbO$_3$ crystals as well as the thicker base crystal. These resonances produce spurious signals which are usually 30-40 dB below the correlation output signal.

For the cleaning and assembly we have applied with small modifications the technique des-cribed by Smith (1975). The silicon is supported on an RTV-pillow to maintain a uniform pressure. In addition to applying plastic coatings to the silicon and the delay line surface to peel off dust particles, it is usually necessary with an additional manual cleaning to rip off the last few stubborn dust particles. The whole assembly can normally be completed within ½ hour.

In order to maintain a controlled gap, we have applied chromium dots of 200-300 nm height rather than ion beam etched posts (INGEBRIGTSEN et al. 1975). The chromium dots of dia-meter 2-3 μm are randomly positioned on the delay line surface with a density of 100-200 dots per mm^2. We have not observed any artifacts which can be related to the use of chromium dots instead of LiNbO$_3$ posts.

A specific problem in the package design is to eliminate the electromagnetic feedthrough between input and output circuitry. It is usually sufficient to suppress this feedthrough by 100 dB. However, this requires both a careful design of the package to suppress coupling through the package itself as well as specific means to reduce the electromagnetic coupling through the LiNbO3 crystal. The latter can be controlled by applying grounded metal shields around the input transducer on the surface of the thin LiNbO3 plate as well as on the interface to the thicker base crystal.

3.   EXPERIMENTAL RESULTS

Several devices using platinumsilicide Schottky diodes with periods ranging from 5 μm to 12.5 μm and pn-diodes of period 12.5 μm have been fabricated. The Schottky diode devices have been operated at center frequencies of 100 MHz and 200 MHz, the pn-diode devices around 100 MHz center frequency. The purpose of the experiments has been to evaluate the fabrication technique, to establish a design model, and to explore performance limits. The results of the experiments carried out so far may be summarized as follows.

   - The fabrication procedure is adequate to control interaction uniformities within 2 dB over interaction lengths of 2-3 cm corresponding to integration times of 6-9 μsec.

   - The configuration with a thin LiNbO3 plate glued onto a thicker base crystal results in spurious levels ∿ 30 dB below the correlation signal over a 20 per cent bandwidth.   .

   - Schottky diodes can be charged with a 2 nsec plate pulse, while pn-diodes require integrated charging with an rf-burst of ∿ 0.5 μsec time duration.

   - Storage times range from 5-50 msec in Schottky diodes and up to 100 sec in pn-diodes. The storage time may vary by a factor of 2-3 over interaction lengths of 2-3 cm for both kinds of diodes. The storage time increases somewhat with the plate pulse amplitude. This is particularly notable for pn-diodes which require a long rf-burst on the plate. The storage time in pn-diodes is frequency dependent with shorter storage times for higher frequencies. The observed voltage and frequency dependence of the storage time in pn-diodes is demonstrated in Figs. 5 and 6.

   - We have observed, for Schottky diodes and pn-diodes, a frequency dependence of the correlation output which is not observed when the same devices are operated in a convolution mode. Subtracting all known frequency dependent factors such as plate pulse duration, diode geometry, and measured frequency dependence of the storage time, a substantial reduction in the correlation efficiency with increasing frequency is observed. With Schottky diodes we have observed a fall off towards higher frequencies proportional with the frequency to the 3rd and 4th power, and with pn-diodes as high as the 6th power. An example of this is shown in Fig. 7 for a pn-diode device. We have no explanation of these observations, but the experiments point towards a deficiency in the charging process. We believe that it is important to understand and maybe find a solution to this problem since it may ultimately lead to the fabrication of more efficient devices.

4.   MODELLING

For the configuration shown in Fig. 8 a per unit area equivalent circuit model has been applied for the analysis. A similar model has recently been proposed by Kino (KINO, G., 1977). The problem may be separated in space-independent fields and currents related to the plate circuit and space dependent fields associated with the acoustic surface wave. With reference to the experimental conditions it is necessary to carry out a large signal analysis for the space uniform plate current during charging, whereas the current associated with the acoustic wave can be treated as a superimposed small signal current.

The per unit area equivalent circuit for the space uniform plate circuit is shown in Fig. 9. Here C_L is the capacitance of the LiNbO3 crystal, C_O the capacitance of the air gap underneath the diodes, and C_S is the resultant capacity between the LiNbO3 crystal and the silicon of the areas in between the diodes. This model assumes that the resistivity of the bulk silicon is negligible and that the surface of the LiNbO3 crystal is an equipotential surface. The first assumption is good since the dominating series resistance that limits the rise time in practice will be the internal resistance of the plate pulse source. The latter assumption relies on the high relative permittivity (∿70) of LiNbO3.

The space variable field from the surface wave is assumed to originate from a polarization charge

$$P = P_O \cos(\omega t - kz) \qquad (4)$$

located at the surface of the delay line. This polarization is equal in magnitude, but of opposite sign to the charge induced on a shorting plane at the surface of the piezoelectric. The polarization charge can be related to the sheet power per unit width of the surface wave beam, S, by (INGEBRIGTSEN 1969)

$$P_O = 2\left(\frac{\Delta v}{v}\right)^{\frac{1}{2}} \frac{(\omega \varepsilon_p S)^{\frac{1}{2}}}{v} \qquad (5)$$

where $(\Delta v/v)$ is the surface wave coupling coefficient (0.024 for YZ-LiNbO3), ε_p the effective surface wave permittivity ($50\varepsilon_0$ for YZ-LiNbO3), and v the surface wave velocity (3486 m/s for YZ-LiNbO3).

The fields originating from this charge partly penetrate into the delay line and partly bridge the air gap and penetrate into the semiconductor. In the analysis we shall neglect the field in the bulk of the semiconductor. This is permitted provided the induced voltage across the air gap and the diodes dominate over the induced voltage within the bulk of the semiconductor. Neglecting the diodes the approximation holds when

$$k\sigma \gg \omega C_0 \qquad (6)$$

where k is the wavenumber, σ the bulk conductivity of the semiconductor, ω the frequency, and C_0 the air gap capacitance per unit area. With an air gap of 200 nm the approximation holds for bulk resistivities much less than 150 Ωcm. Neglecting the fields in the semiconductor has as consequence that the model does not account for acoustoelectric loss. However, when the losses are small, it does not affect the per unit area equivalent circuit which aims at explaining local effects. The losses can therefore phenomenologically be included by letting the polarization charge amplitude P_0 be space dependent.

With these assumptions the equivalent circuit for the space variable terms are shown in Fig. 10. The only modification of the equivalent circuit for the space uniform terms is to add a shunt capacitance C_p to account for the fields from the polarization charge which penetrates into the delay line. This capacitance is per unit area

$$C_p = k \, \varepsilon_p \qquad (7)$$

where k is the wavenumber and ε_p (= $50 \, \varepsilon_0$ for YZ-LiNbO3) is an effective permittivity for the piezoelectric.

Assuming that the diode capacitance is constant and voltage independent and that the current voltage characteristic of the diodes follows an ideal exponential law, i.e.

$$I = I_0(e^{\beta V_d} - 1) + C_d \frac{dV_d}{dt} \, , \qquad (8)$$

where I_0 is the diode saturation current, $\beta = e/kT$, V_d the diode voltage, and C_d the diode capacitance, we can find exact analytic solutions which cover the recording and storage of the space dependent diode charge.

Assuming that the diode current resulting from the surface wave polarization charge is small compared to the large signal current from the plate pulse voltage $V_p(t)$, the space varying charge stored on the diodes is

$$Q_\sim(kz,\omega) = [P_0\gamma(\omega)] \times \frac{e^{-\beta Q_0(T_p)/C_2}}{\tau_2} \int_0^{T_p} \cos(\omega\tau - kz)e^{\alpha\beta V_p(\tau)} \cdot e^{\beta Q_0\left(\frac{1}{C_2} - \frac{1}{C_1}\right)} d\tau \qquad (9)$$

where T_p is the time duration of the plate pulse. Q_0 is the magnitude of the space uniform stored charge resulting from the plate pulse, defined through

$$e^{\frac{\beta Q_0(t)}{C_1}} = 1 + \frac{1}{\tau_1} \int_0^t e^{\alpha\beta V_p(\tau)} d\tau \, . \qquad (10)$$

For the diode orientation shown in Fig. 9, this charge is actually negative. The other quantities in Eqs. (9) and (10) are defined by

$$C_1 = C_d + \frac{C_0(C_L + C_S)}{C_0 + C_L + C_S} \, , \qquad (11)$$

which is the resultant capacitance seen from the diode terminals in the equivalent circuit of Fig. 9.

$$C_2 = C_d + \frac{C_0(C_L + C_S + C_p)}{C_0 + C_L + C_S + C_p} \, , \qquad (12)$$

which is the similar capacitance for the equivalent circuit in Fig. 10. Note that since C_p is proportional with frequency, C_2 is frequency dependent.

$$\gamma = \frac{C_0}{C_0 + C_L + C_S + C_p} \, . \qquad (13)$$

This means that the first term in Eq. (10) $[P_O\gamma]$ is equal to the charge density induced on an ideal shorting plane located at the surface of the semiconductor. $[P_O\gamma]$ is therefore the maximum charge that can be stored on the diodes.

$$\alpha = \frac{C_O C_L}{C_1(C_O+C_L+C_d)} \quad . \tag{14}$$

Thus αV_p is the voltage across the diodes, if the conduction current through the diodes is neglected.

The time constants τ_1 and τ_2 are respectively the small signal time constants at zero voltage seen across the diode terminals of the equivalent circuits of Figs. 9 and 10. They are defined by

$$\tau_1 = C_1/\beta I_O \qquad \text{and} \qquad \tau_2 = C_2/\beta I_O \tag{15}$$

Since C_2 is frequency dependent, τ_2 will also be.

A further simplification of Eq. (10) is obtained when the plate voltage V_p is a pulse of constant amplitude and time duration T_p much smaller than the rf-period. Then

$$Q_\nu(kz,\omega) = [P_O\gamma] \times \left[\frac{e^{\beta Q_O/C_2} - 1}{e^{\beta Q_O/C_2}}\right] \cos kz \quad , \tag{16}$$

where

$$e^{\beta Q_O/C_2} = \left[1 + T_p/(\tau_1 e^{-\alpha\beta V_p})\right]^{(C_1/C_2)} \quad . \tag{17}$$

Since $\tau_1 e^{-\alpha\beta V_p}$ is the differential time constant at the initial diode bias αV_p, we conclude from Eq. (16) that a complete charging is obtained provided this time constant is much smaller than the charging time, and that the capacitance ratio (C_1/C_2) is not too small. When this is the case, the frequency dependence of C_2 is insignificant and the only frequency dependence in Q_ν arises from the frequency dependence of the factor γ. γ is constant for frequencies

$$\omega < \frac{v(C_L + C_s + C_O)}{\varepsilon_p} \quad . \tag{18}$$

For higher frequencies it is inversely proportional with the frequency. Usually this characteristic frequency is of the order of 50-100 MHz.

After the plate pulse is turned off, the stored charge decays with time as

$$\frac{Q_\nu(t)}{Q_\nu(o)} = \frac{e^{\beta Q_O/C_2 - t/\tau_2}}{\left[1 + (e^{\beta Q_O/C_1} - 1)e^{-t/\tau_1}\right]^{C_1/C_2}} \tag{19}$$

where Q_O is the uniform charge on the diodes at $t = 0$. Since τ_2 is larger than τ_1 there is no decay of this charge before the uniform charge has decayed $[\exp(\beta Q_O/C_1) \gtrsim \exp(t/\tau_1)]$. Thereafter the space dependent charge decays with the time constant τ_2. Since τ_2 is a function of frequency, the decay will also be. Since τ_2 increases with the frequency this model predicts a slower decay for higher frequencies. However, this frequency dependence is rather small, and τ_2 will usually be constant for frequencies above 100 MHz.

Figure 11 shows calculations of the decay for various values of the parameter $\exp\{\beta Q_O/C_1\}$. Increasing values of this parameter correspond to increasing plate voltages V_p. Thus we observe that the decay is depending on the plate impulse V_p. This is consistent with observations of Defranoult et al. (1976) and with our observations (see Fig. 5). Also shown in Fig. 11 is the frequency dependence of the decay for the parameter value of 100. A very small frequency dependence is predicted for frequencies between 50 MHz and 200 MHz. This is not consistent with the experimental results shown in Fig. 5.

The read out is obtained as a nonlinear mixing of the time constant space varying stored diode charge and the time and space dependent displacement current through the diodes resulting from the read out signal. This provides an output signal which is proportional with the correlation integral over the interaction length of the stored charge with the read out signal. Principally this mixing arises in the nonlinear diode capacitances. The uniform space independent charge trapped on the diodes through the charging is divided between the diode capacitance (C_d) and the capacitance of the air gap (C_O) in series with the delay line plate capacitance (C_L) (see Fig. 9). This results in a reverse bias on the diodes which determines the operating point on the diode C(V) curve for the nonlinear mixing process.

To analyze the read out process it is assumed that the voltage across the diodes is related to the charge by

$$V_d(Q_t) \simeq V_d(Q_o) + \frac{Q_t - Q_o}{C_d} + \kappa (Q_t - Q_o)^2 \; , \tag{20}$$

where V_d is the diode voltage, Q_t the total diode charge, Q_o the space uniform bias charge, C_d the differential capacitance at $V_d(Q_o)$. The nonlinear coefficient κ is then defined by

$$\kappa = \frac{1}{2} \frac{d^2 V_d}{dQ^2} \bigg|_{Q=Q_o} = - \frac{1}{2} \frac{1}{C_d^3} \cdot \frac{dC_d}{dV_d} \bigg|_{Q=Q_o} \tag{21}$$

Since we can neglect the conduction current through the reverse biased diodes, the polarization charge P_1 from the read out signal will give rise to an electric displacement through the diodes

$$P_1 \gamma \frac{C_d}{C_2} \cos(\omega t - kz) \tag{22}$$

Correspondingly the stored space variable diode charge is responsible for an electric displacement through the diodes which is

$$Q_\sim \frac{C_d}{C_2} \cos kz \tag{23}$$

The sum of these two terms and the space uniform charge Q_o constitutes the total diode charge Q_t. The nonlinear term in Eq. (20) gives rise to a space independent voltage at the frequency ω across the diodes which is proportional to the product of the amplitudes of the space varying terms. With the plate circuit externally shorted a short circuit current per unit area results which is

$$i_s = j\omega\kappa\alpha\gamma \, C_d \left(\frac{C_d}{C_2}\right)^2 P_1 Q_\sim \cos \omega t \tag{24}$$

To account for acoustoelectric loss it may be assumed that P_1 and $Q_\sim$ decay exponentially with increasing z with the rate $e^{-\alpha z}$. The total short circuit current from an interaction region of width W and length L is then

$$I_s = i_s A \frac{1 - e^{-2\alpha L}}{2\alpha L} \tag{25}$$

where the plate area $A = WL$.

With an external impedance Z_L connected to the plate terminals a voltage appears across Z_L which is

$$V_{out} = I_s \cdot \frac{Z_L}{1 + j\omega C_i Z_L} \tag{26}$$

where

$$C_i = A(1 + \alpha \frac{C_d}{C_s}) \frac{C_L C_s}{C_L + C_s} \tag{27}$$

is the internal capacitance of the plate circuit.

The plate circuit may be tuned in series or in parallel with an inductive load. The quality factor Q_L of this load is limited by the desired bandwidth B to $Q_L \leq 1/B$. The available output power from the plate is then

$$P_{out} = \frac{1}{2} |I_s|^2 \frac{Q_L}{\omega C_i} \tag{28}$$

The results obtained above may be applied to discuss the frequency dependence of the output signal.

Since $C_p(\omega)$ usually dominates in the denominator of γ (Eq. (13)) for frequencies above 50-100 MHz, γ will be inversely proportional to the frequency. For the same reason C_2 (Eq. (12)) will be essentially frequency independent. According to Eqs. (5) and (16) $P \propto \omega^{\frac{1}{2}}$ and $Q_\sim \propto P_o \gamma \propto \omega^{-\frac{1}{2}}$, the short circuit current per unit area, i_s, is frequency independent.

Consistent with the approximations the spatially varying rf-current in the semiconductor is determined by the capacitance of the air gap and the diodes (see Eq. (6)). Thus the spatially varying current per unit area $i_A = j\omega P_{0,1}\gamma(C_0+C_S)/C_0$. The sheet conductance of the bulk semiconductor is $G = k\sigma$ where σ is the bulk conductivity, and k the wavenumber. With the assumed frequency dependence of P and γ, it follows that the loss per unit area $|i_A|^2/G$ and hence the attenuation coefficient α, is frequency independent. In conclusion, the external short circuit current I_S given by Eq. (25) is frequency independent.

Assuming that the same read out process may be applied to describe convolver operation where two counter propagating waves are mixed to produce a convolution signal at the double frequency, Eq. (24) still applies with the modification of a factor of 2 in frequency and that Q_ν is replaced by $P_2\gamma$ where P_2 is the polarization charge corresponding to the counter-propagating wave. This, however, leads to the same conclusion as above, namely that the short circuit convolution current is frequency independent.

To summarize we conclude that although this simple model describes the essential principles of operation of the diode memory correlator it is inadequate to describe the frequency dependence of the correlation efficiency that we have observed experimentally. Since the observations give a frequency dependence in convolver operation in agreement with the theory, the model for the read out process is believed to be correct. One should therefore look for explanations in the recording or in the storage process.

5.   CONCLUSION

With the fabrication technique described one can fabricate diode memory correlators with a dynamic range limited by spurious signals to 30 dB. Improvements require that acoustic plate resonances can be further suppressed. The intrinsic conversion loss is of the order of -70 to -80 dBm depending on the silicon resistivity and the coupling strength in the interaction region. Consistent with earlier observations devices using Schottky diodes can be charged with nanosecond pulses, thus accommodating signals with bandwidths in excess of 100 MHz. The observed storage times are of the order of 10 milliseconds. Devices using pn-diodes are slower and require integrated charging over several hundred nanosecond. These devices can on the other hand store signals for as ng as 100 seconds. Even permanent storage (for several hours) has occasionally been observed at room temperature in devices which have been carefully protected from external illumination.

The simple model presented has demonstrated its ability to explain important features such as plate pulse requirements, storage times, and the essential features of the correlation read out process. Experimentally we have observed with Schottky diode and pn-diode devices a strong reduction in the correlation efficiency with increasing frequency through the design pass bands. Similar frequency dependence is not observed when the devices are operated as convolvers. Since the read out processes in correlation and convolution probably are the same, it is concluded that the observed frequency dependence in correlation must be related to the recording and storage process. It is important to solve this problem since it may ultimately lead to fabrication of more efficient devices.

6.   REFERENCES

BORDEN, P. and KINO, G.S., 1977, "Input Correlation with the ASW Storage Correlator", Electronics Letters, 13, pp. 470-471.

BORDEN, P.G., JOLY, R. and KINO, G.S., 1976, "Correlation with the Storage Convolver", 1976 Ultrasonics Symposium Proceedings, IEEE Cat. no. 76 CH1120-5SU.

DEFRANOULT, Ph., GAUTIER, H., MAERFELD, C. and TOURNOIS, P., 1976, "PN-diode Memory Correlator", 1976 Ultrasonics Symposium Proceedings, pp. 336-347, IEEE Catalogue Number 76 CH 1120-SU.

INGEBRIGTSEN, K.A., 1969, "Surface Waves in Piezoelectrics", J. Appl. Phys. 40, pp. 2681-2686.

INGEBRIGTSEN, K.A., COHEN, R.A. and MOUNTAIN, R.W., 1975, "A Schottky-Diode Acoustic Memory and Correlator", Appl. Phys. Lett., 26, pp. 596-598.

INGEBRIGTSEN, K.A., 1976, "The Schottky Diode Acoustoelectric Memory and Correlator - A Novel Programmable Signal Processor", Proc. IEEE, May 1976.

KINO, G.S., 1977, private communication.

RALSTON, Richard W., 1975, "Stable CW Operation of Gap Coupled Silicon on Sapphire to LiNbO3 Acoustoelectric Amplifiers", 1975 Ultrasonics Symposium Proceedings, pp. 217-222, IEEE Catalogue Number 75 CHO 994-4SU.

RALSTON, R.W., HURLBURT, D.H., LEONBURGER, F.J., CAFARELLA, J.H. and STERN, E., 1977, "Gap-Coupled Schottky Diode/LiNbO3 Acoustoelectric Integrating Correlator", paper presented at the 1977 Device Research Conference, Cornell University, June 27-29, 1977.

REIBLE, S.A., CAFARELLA, J.H., RALSTON, R.W. and STERN, E., 1976, "Convolvers for DPSK Demodulation of Spread Spectrum Signals", 1976 Ultrasonics Symposium Proceedings, pp. 451-455, IEEE Cat. no. 76 CH 1120-5SU.

SMITH, Henry J.,1975, "Techniques for Making Gap Coupled Acoustoelectric Devices", 1975 Ultrasonics Symposium Proceedings, pp. 238-240, IEEE Catalogue Number 75 CHO 994-4SU.

STERN, E. and WILLIAMSON, R.C., 1974, "New Adaptive Signal Processing Concept", Electronics Letters, $\underline{10}$, pp. 58-59.

TAKADA, S., HOH, K., HAYAKAWA, H. and TOKUMARU, Y., 1976, "Surface Wave Memory-Correlator Using Light Pulse from a GaAs Laser Diode", 1976 Ultrasonics Symposium Proceedings, IEEE Catalogue no. 76 CH1120-5SU.
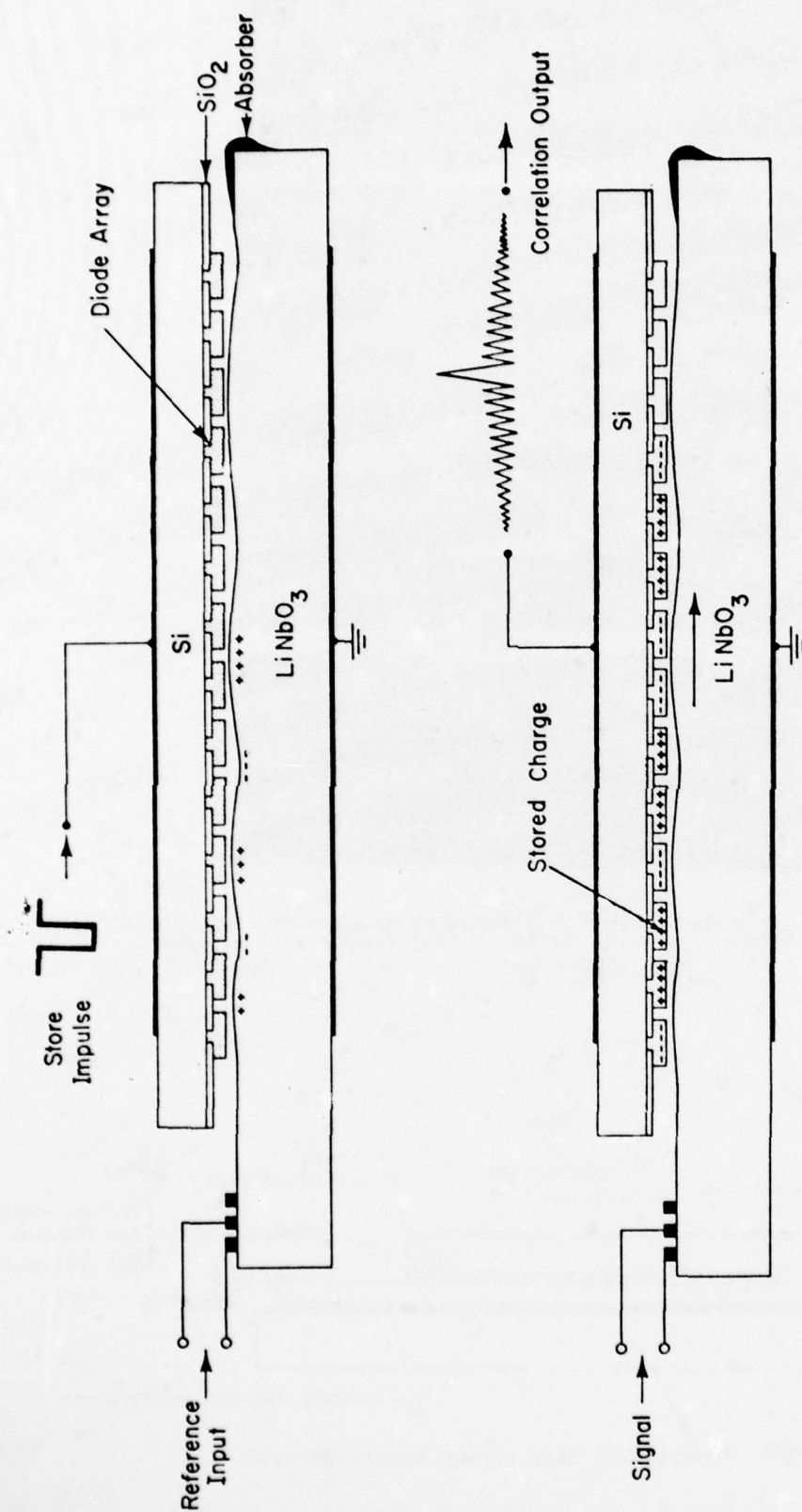
Fig. 1. Principal operation of gap coupled diode memory correlator in recording (upper part of the figure) and in correlation (lower part of the figure).
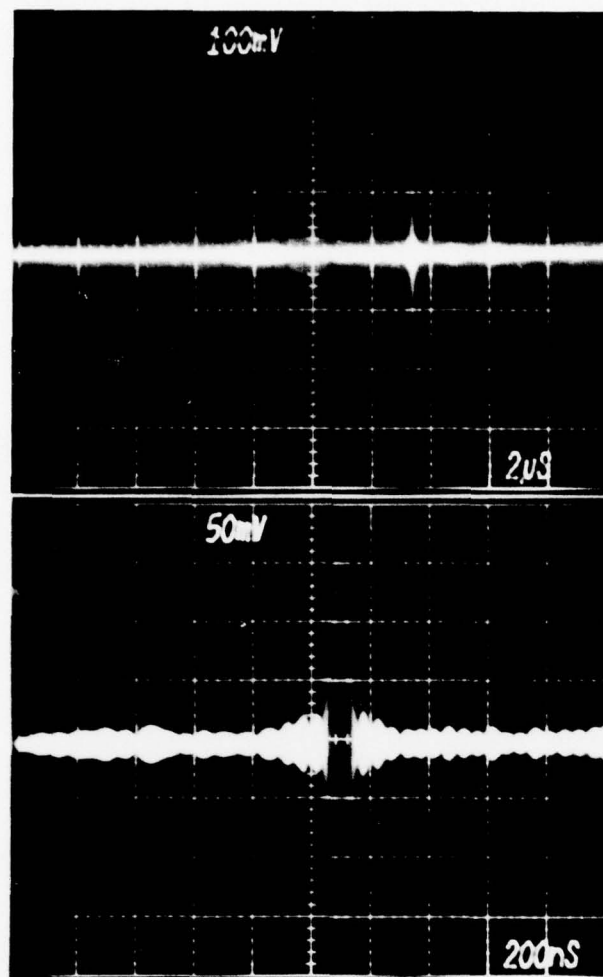
Fig. 2.   Chirp correlation in a Schottky-diode memory correlator.
Center frequency 200 MHz, chirp frequency excursion
30 MHz, chirp time excursion 5 usec. Storage time 3
millisecond.
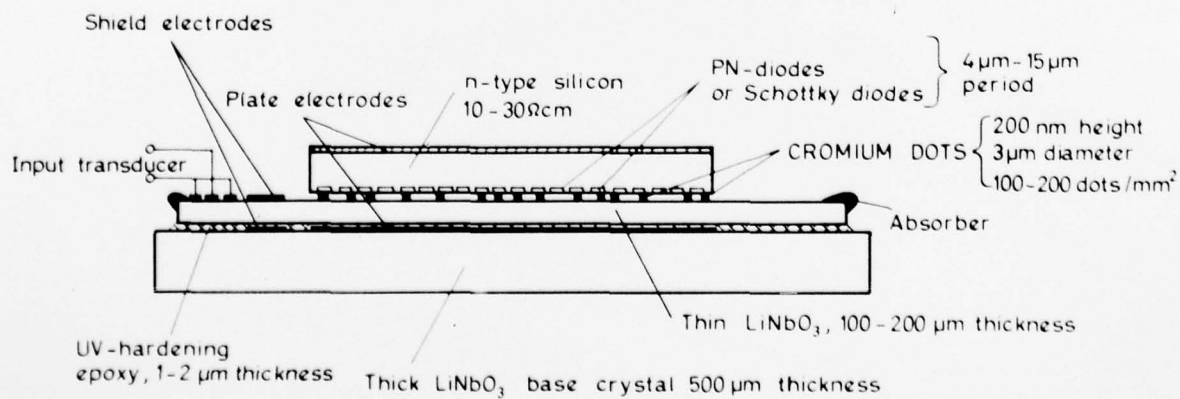
Fig. 3.   Schematic of the experimental configuration.
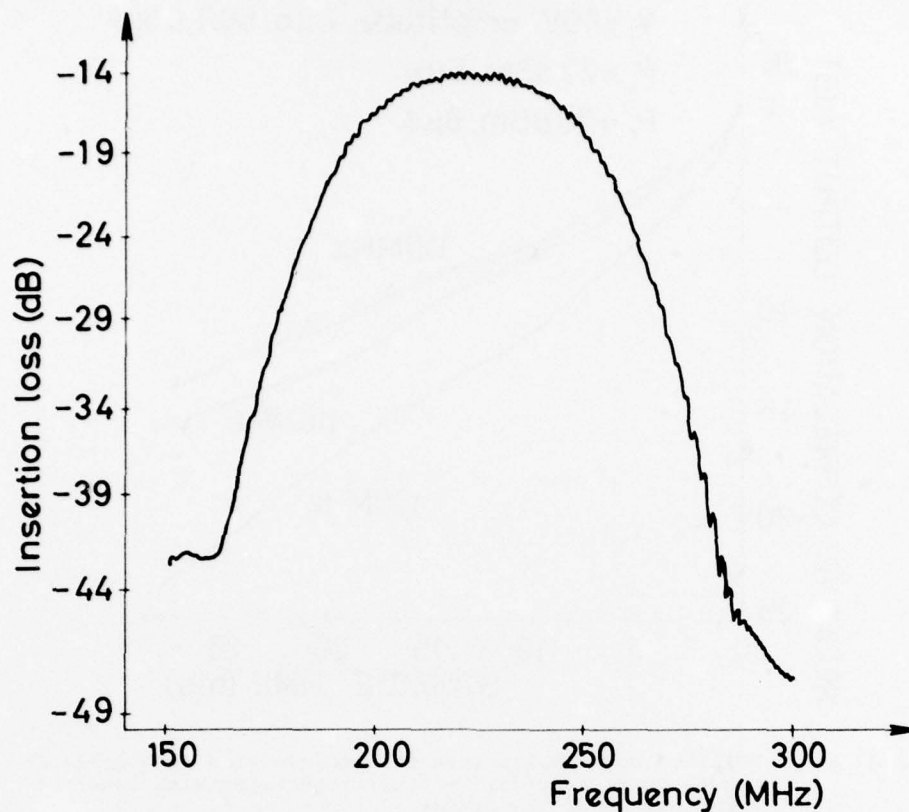
Fig. 4.   Frequency response of a 12 μsec surface wave delay line
on a 150 μm thick LiNbO₃ crystal. The ripple is caused
by reflections off an auxiliary transducer.
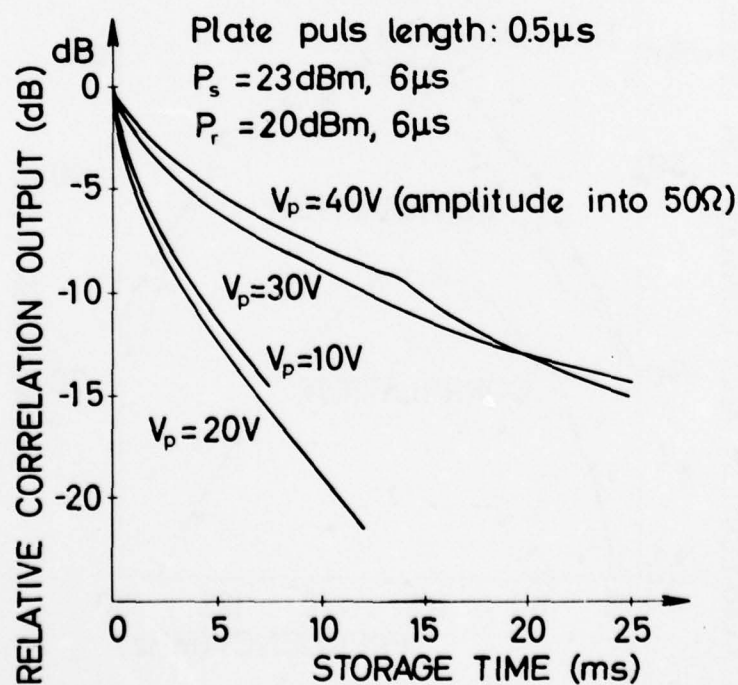
Fig. 5.   Correlation loss vs. storage time for a 100 MHz pn-diode
memory correlator with various plate burst amplitudes.
A constant illumination was applied to reduce storage
times from ∿ 10 sec to the millisecond range.
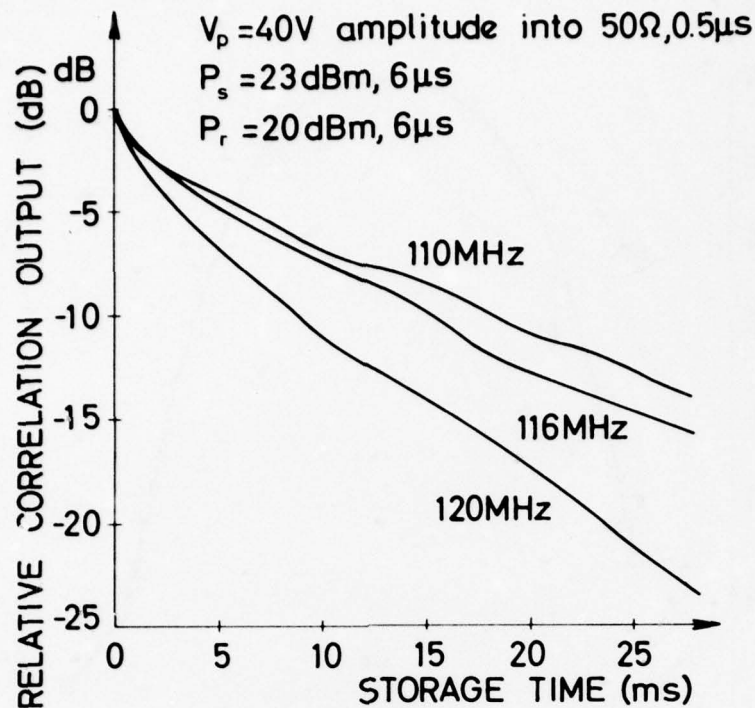
Fig. 6. Relative correlation loss vs. storage time in a pn-diode memory correlator for various frequencies with constant illumination to the diodes.
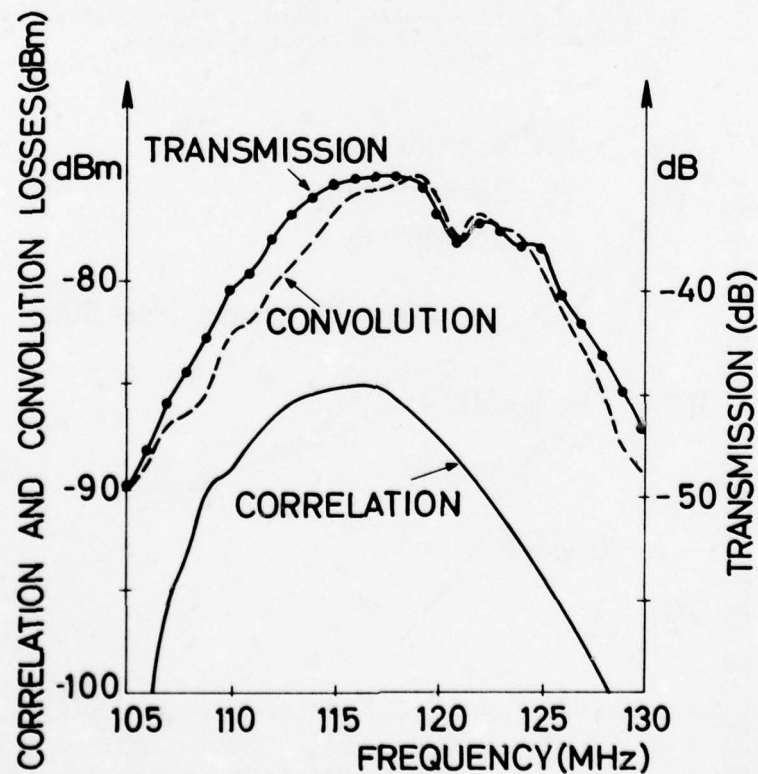
Fig. 7. Transmission loss, convolution loss, and correlation loss vs. frequency for a pn-diode device. The high transmission loss and convolution loss and the dip in the frequency response at 121 MHz is caused by a default in one of the transducers. The good transducer was used for the correlation measurement.

CORRELATOR CONFIGURATION
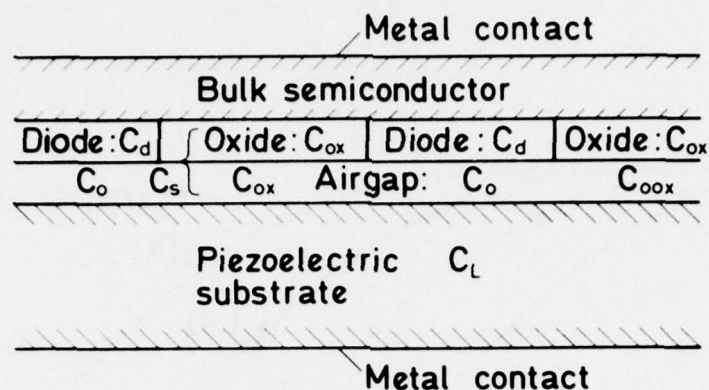
Fig. 8.   Schematic of the configuration used for modelling.
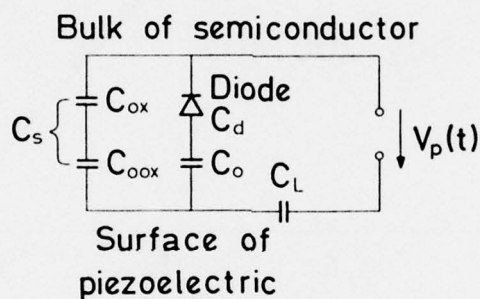
Fig. 9.   Equivalent circuit applied to analyze the space uniform fields.
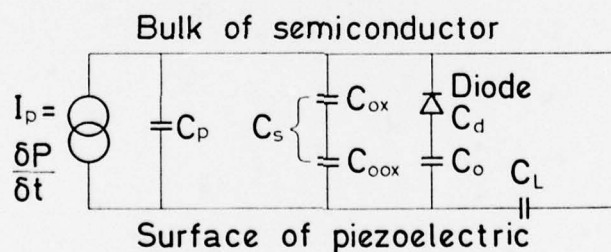
Fig. 10.   Equivalent circuit applied to analyze the wave generated space varying fields.
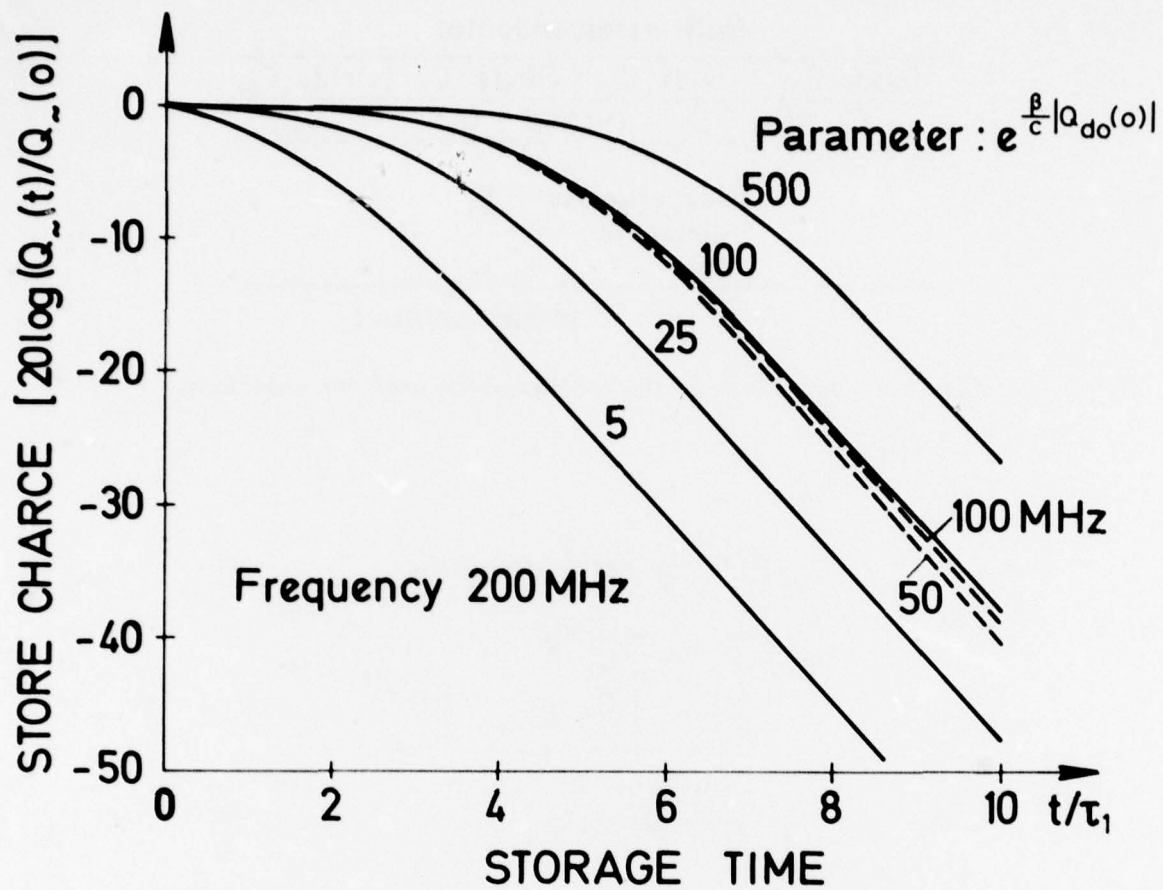
Fig. 11. Calculated decay showing the plate pulse dependence and frequency dependence of the storage predicted by the model.

CONVOLUTION ET CORRELATION A MEMOIRE A L'AIDE DE COMPOSANTS ACOUSTIQUES A ONDES DE SURFACE[*]

H. Gautier, C. Maerfeld et P. Tournois
Thomson-CSF - Division des Activités Sous-Marines
Chemin des Travails - 06802 Cagnes-sur-Mer (France)

Les systèmes de traitement du signal comportent en général un premier étage permettant une mise en forme des signaux suivi d'un module servant à l'identification, la comparaison ou la mesure des signaux reçus. Pour ce faire, les opérations qu'il est souvent besoin de réaliser sont le filtrage linéaire et le produit de convolution ou celui de corrélation, auxquelles vient s'ajouter la notion de mémoire.

L'opération mathématique de base est le produit de convolution. En effet, pour tout filtre linéaire causal on peut définir une réponse impulsionnelle. Soit G(t) celle-ci, le signal de sortie du filtre est alors la convolution du signal d'entrée avec G (Voir Fig. 1a). Dans un filtre classique, la fonction G est câblée, donc figée, et le système fonctionne de manière totalement asynchrone.

De façon plus générale, on peut envisager un dispositif à deux entrées et une sortie capable de générer le produit de convolution des deux signaux présents à l'entrée. Une solution simple qui utilise des ondes acoustiques de surface contradirectives comme support de l'information à traiter, a été mise au point. Elle fait l'objet de la section 1 de cette présentation. Du fait de la contradirectivité des ondes, nous verrons que le signal de convolution subit une compression dans le temps d'un facteur 2 (voir Fig. 1b). Notons que pour utiliser le système en corrélateur il faut retourner dans le temps l'un des signaux d'entrée ; en outre le bon fonctionnement de l'appareil nécessite en général le synchronisme de G et F.

Pour s'affranchir de ces contraintes, un dispositif acoustique du même type a été réalisé . Il présente une mémoire interne qui permet de conserver pendant un certain temps un signal, par exemple G. On montre alors que le système se comporte comme un filtre linéaire de réponse impulsionnelle G(t) (convoluteur) ou G(-t) (corrélateur) (Voir Fig. 1c). La fonction G peut être changée à volonté , le filtre est donc programmable. Ce type de composant est actuellement au stade d'étude en laboratoire et nous décrirons dans la Section 2 ses caractéristiques et les spécifications que l'on peut espérer d'ici quelques années.

Les composants acoustiques des types de ceux que nous présentons ici sont bien adaptés au traitement des signaux en fréquence intermédiaire utilisés dans le domaine de l'Avionique. Ces composants peuvent en effet traiter directement des signaux analogiques FI de largeur de bande et de durée allant jusqu'à 100 MHz et 20 μs environ respectivement. Leur vitesse de calcul et leur encombrement rendent leur emploi intéressant quand la quantité d'information à traiter devient importante.

1. - LE CONVOLUTEUR ACOUSTIQUE.

1.1. Principes.

Afin de générer le produit de convolution de deux signaux, il est nécessaire de disposer dans le système d'un phénomène non linéaire. Ce mécanisme peut être simplement la non linéarité intrinsèque du matériau acoustique lui-même. Quate et Thompson [1] ont fait les premières études de cet effet avec des ondes de volume ; Svaasand [2], Luukkala et Kino [3] ont étudié les interactions d'ondes de surface.

Pour décrire le fonctionnement du système, nous prenons le cas d'un simple convoluteur piézo-électrique à onde de surface, qui est schématisé sur la Fig. 2. Il est constitué d'un cristal piézoélectrique poli sur la surface duquel ont été déposés deux transducteurs interdigités. Ces transducteurs sont alimentés par des signaux de même pulsation ω et d'amplitude F(t) et G(t) respectivement. Les deux ondes acoustiques contradirectives ainsi émises seront de la forme $F(t - z/v) e^{j(\omega t - kz)}$ et $G(t + z/v) e^{j(\omega t + kz)}$ où v est la vitesse de propagation et k = ω/v le nombre d'onde.

Du fait des propriétés non linéaires du substrat, des termes de potentiel du deuxième ordre sont engendrés. En particulier le terme à fréquence 2ω correspond au produit des deux formes d'ondes précédemment citées. Il est donc spatialement uniforme et il peut être détecté au moyen de deux électrodes planes déposées sur les surfaces supérieure et inférieure du substrat. Ces électrodes recueillent la somme de tous les courants non linéaires générés sous elles ; le signal recueilli en sortie est donc de la forme : $H(t) e^{j2\omega t}$ avec :

$$H(t) = C \int_{-L/2}^{L/2} F(t - \frac{z}{v}) \, G(t + \frac{z}{v}) \, dz \qquad (1)$$

Si les signaux F et G sont de durée θ au plus égale à T = v/L et si leur retard relatif est au plus égal à (T - θ)/2, alors les bornes de l'intégrale (1) peuvent être amenées à ∞, et au moyen d'un changement de variable τ = t - z/v, on peut écrire :

$$H(t) = D \int F(\tau) \, G(2t - \tau) \, d\tau \qquad (2)$$

H(t) est ainsi le vrai produit de convolution de F et G comprimés dans le temps dans un rapport 2 à cause de la contradirectivité des ondes incidentes. D est lié au rendement énergétique du système. Ce dernier est relativement faible quand on utilise le système tel qu'il est représenté sur la Fig. 2. Afin d'améliorer la dynamique nous avons mis au point un dispositif qui utilise des compresseurs de faisceau acoustique [4] et qui permet d'augmenter la densité de puissance acoustique avant mélange et par là le rendement. Ce composant sera détaillé aux paragraphes qui suivent.

Dans le même but il est intéressant d'utiliser la plus forte non linéarité de l'interaction des champs electriques associés aux ondes acoustiques dans des milieux piézoélectriques avec les porteurs de charge d'un milieu semiconducteur. Plusieurs structures ont été étudiées ; elles sont représentées sur la Fig. 3 avec références aux principaux travaux effectués.

La structure type 3a est composée d'une plaquette de semiconducteur maintenue à une faible distance du piézoélectrique [5-8]. Le film d'air présent est nécessaire pour éviter toute charge mécanique de la surface du piézoélectrique. Elle offre l'avantage de pouvoir optimiser séparément les deux milieux et d'obtenir ainsi les plus forts rendements enregistrés à ce jour. Elle présente néanmoins l'inconvénient de ne pas être monolithique et de nécessiter en particulier la présence d'un film d'air de l'ordre d'une fraction de micron. Le couplage avec le semiconducteur n'est stable et peu sensible à la température qu'au prix d'une assez grande complexité mécanique [8]. Pour s'affranchir du film d'air, des structures monolithiques ont été expérimentées qui utilisent soit un semiconducteur piézoélectrique comme le CdS ou le GaAs [9] (voir Fig. 3b), soit une couche mince piézoélectrique déposée sur un substrat semiconducteur par exemple le couple ZnO/Si [10] (voir Fig. 3c). Enfin, il faut citer une dernière géométrie qui est celle d'un convoluteur hybride composé d'une ligne acoustique à prises multiples où le mélange non linéaire a lieu dans un réseau extérieur, généralement intégré, de diodes semiconductrices[11] (voir Fig. 3d).

Pour tout convoluteur, la puissance du pic de convolution P_{CONV} varie proportionnellement au produit des puissances des signaux électriques d'entrée $P_1 P_2$ tant que le système fonctionne dans son régime bilinéaire. On définit alors le facteur de mérite F suivant :

$$F \ (dBm) = 20 \ \log \left\{ P_{CONV} \ (mW) \ / \ P_1 \ (mW) \ \times \ P_2 \ (mW) \right\}$$

Pour comparer utilement les différents dispositifs présentés précédemment, il convient de préciser les niveaux de puissances maximum admissibles soit pour ne pas détériorer les transducteurs, soit pour conserver au système sa bilinéarité. Ceci revient à préciser le niveau maximum du signal de convolution disponible. Enfin, les rendements obtenus dépendent aussi du nombre de points (produit largeur de bande par durée des signaux) qu'il est possible de traiter.

Le tableau I compare les différents dispositifs précités sur la base de leur rendement F, du niveau maximum de puissance admissible ou disponible et du nombre de points actuellement démontré ou que l'on peut envisager de façon réaliste.

Il ressort de cette comparaison que les systèmes aux rendements les plus élevés présentent des niveaux de saturation faibles et qu'à capacité de traitement identique, leur dynamique n'est pas supérieure à celle des dispositifs moints efficaces. Les dispositifs monolithiques (ZnO, GaAs) ont actuellement une bande passante limitée à 10 à 20 MHz. Aussi deux systèmes seulement permettent réellement de traiter des signaux de produit bande-temps supérieur à 100 ou 200 : ce sont les convoluteurs à milieu semiconducteur rapporté et les convoluteurs piézoélectriques à compresseur de faisceau. Nous allons maintenant détailler une réalisation particulière de ce dernier type de convoluteur.

- TABLE I -

Comparaison de divers types de convoluteurs acoustiques

Type de convoluteur	Facteur de bilinéarité F (dBm)	Nombre de points = B.T	Puissance maximum des signaux (dBm)		Référence
			d'entrée	de sortie	
Piézoélectrique simple	-85	> 100	30	-25	[3]
Piézoélectrique à compresseur de faisceau	-70	> 100	30	-10	[4]
A semiconducteur et film d'air	-45 / -65	~ 30 / > 100	27 / 30	0 / -5	[7 - 8]
ZnO sur Si	-58	≤ 100	25	-8	[10]
A GaAs	-50 / -60	~ 30 / ≤ 100	25 / 25	0 / -10	[11]
A réseau de diodes extérieur	-50 / -66	~ 30 / ≤ 150	30 / 30	30 / -6	[11]

1.2. Convoluteur piézoélectrique à compresseur de faisceau.

1.2.1. Géométrie.

La géométrie du dispositif est plane et elle est schématiquement représentée sur la Fig. 4. Sur un substrat piézoélectrique poli en LiNbO$_3$ sont déposés deux transducteurs identiques qui permettent la génération des ondes acoustiques contradirectives à la pulsation ω. Des coupleurs à bandes métalliques multiples permettent un transfert quasi total de l'énergie incidente et une compression de la largeur de faisceau acoustique dans un rapport 10 à 20. Ceci entraîne donc une augmentation dans le même rapport de la densité de puissance acoustique et donc du rendement non linéaire du système. La largeur des faisceaux comprimés étant de l'ordre de quelques longueurs d'onde acoustique, il est nécessaire de guider l'énergie pour éviter tout phénomène de diffraction excessive. Ce guidage est simplement obtenu grâce à la présence de l'électrode métallique de sortie qui permet de recueillir le signal de convolution. Pour conserver la nature planar de la structure et la simplicité de réalisation qui en découle, une double électrode de masse est déposée sur la surface du substrat.

Le système complet comporte en outre des circuits électriques d'accord des transducteurs d'entrée et en général un transformateur adaptateur d'impédance en sortie. Un exemple de réalisation d'un tel dispositif est donné sur la Fig. 5.

1.2.2. Caractéristiques techniques du convoluteur.

Le tableau II rassemble les principales caractéristiques électriques des dispositifs réalisés à ce jour comme celui de la Fig. 5, ainsi que les limites que l'on doit pouvoir atteindre avec ce type de technologie.

La vitesse de propagation des ondes de surface est de l'ordre de 3500 m/s sur $LiNbO_3$. Aussi avec des cristaux piézoélectriques de longueur hors tout d'environ 50 mm, la durée maximum des signaux que l'on peut traiter de manière synchrone est de 12 µs. Des substrats de 100 mm sont disponibles et on peut donc envisager atteindre des temps de traitement de 20 µs. La fréquence d'entrée pourra passer de 156 MHz à 300 MHz. Ceci nécessitera une précision sur les photogravures d'environ 1 µm.

Le facteur de bilinéarité F défini au paragraphe précédent, est de -72 dBm, et aucun phénomène de saturation n'apparaît jusqu'à des puissances crête d'entrée de 1 W. Au-delà les transducteurs risquent d'être endommagés.(Ils ne supportent en général pas plus de 100 mW de puissance moyenne.) Pour un dispositif de 50 MHz de bande, la dynamique de sortie du système atteint donc plus de 80 dB.

Un autre paramètre important est la précision avec laquelle s'effectue l'intégrale de convolution, c'est-à-dire l'uniformité du traitement sur la durée de celui-ci. On mesure ce paramètre en relevant le signal de convolution d'une impulsion brève avec une impulsion longue. On obtient une représentation directe de l'uniformité. Le résultat est environ +0,2 dB (voir Fig. 6).

La figure de convolution est en principe indépendante de la température car les signaux d'entrée subissent les mêmes variations. Les essais faits entre -20°C et +50°C ont montré qu'aucune déformation du signal de convolution ne se produit. On constate seulement une variation du temps d'arrivée due à la variation du temps de retard moyen du dispositif (de l'ordre de 50 ppm par °C).

- TABLE II -

Caractéristiques techniques du convoluteur piézoélectrique

	Démontré	Limite
Temps de traitement (µs)	12	20
Bande à 3 dB à l'entrée (MHz)	50	100
Fréquence des signaux d'entrée (MHz)	156	300
Facteur de bilinéarité F (dBm)	-72	-72
Puissance maximum d'entrée (W)	1	1
Suppression convolution trajet double (= 2R + A, dB)	40	60
Uniformité de traitement (dB)	+0,2	+0,2
T.O.S.	4	2

1.2.3. Signaux parasites.

Pour terminer la présentation des caractéristiques du système, il est important de décrire les différents signaux parasites. Selon l'utilisation particulière du convoluteur, il sera alors possible d'évaluer la dynamique utile des signaux que l'on peut traiter.

Comme dans tout composant à onde acoustique de surface, les principaux parasites sont dus aux réflexions sur les transducteurs. Chaque onde incidente peut, après réflexion, se combiner avec elle-même ou avec l'autre signal réfléchi donnant ainsi naissance aux signaux parasites dits "d'auto-convolution" et de "convolution par double trajet" respectivement.

Le parasite de "convolution double trajet" est une réplique exacte du signal principal de convolution retardée par rapport à celui du temps de propagation entre transducteurs. Dans un mode de fonctionnement synchrone, ce parasite peut donc être entièrement discriminé par une fenêtre temporelle.

Un tel filtrage ne peut pas être opéré sur le signal parasite d'autoconvolution. Néanmoins il convient de noter qu'un signal ne pourra se convoler avec lui-même que si sa durée est au moins égale au double du retard existant entre transducteur et guide (environ 2 µs). En outre, les signaux utilisés sont généralement codés et dans le cas de codes non symétriques ou aléatoires, de durée θ et de largeur de bande β, le niveau de ce parasite est réduit d'environ $10 \cdot \log(\beta\theta)$ dB.

Dans le tableau III, nous avons résumé les expressions pour les niveaux des différents signaux présents à la sortie d'un convoluteur. Le coefficient de réflexion acoustique sur les transducteurs est dénoté par R, il varie entre 10 et 20 dB pour un transducteur simple selon sa désadaptation. Pour diminuer l'effet des réflexions on aura intérêt à utiliser par exemple des transducteurs unidirectionnels pour lesquels R peut dépasser 30 dB.

L'évaluation de la dynamique du système dépend donc de l'utilisation. On peut néanmoins considérer qu'elle est au moins de l'ordre de 40 dB à 50 dB avec les composants actuels, ce qui est suffisant dans un certain nombre d'applications.

- TABLE III -

Différents signaux d'un convoluteur piézoélec-
trique de capacité maximum B x T

- Signaux d'entrée : durée θ, bande β, puissance P_1 (mW) et P_2 (mW)

- Convolution principale : $10 \log(P_1 P_2) - F - 20 \log \theta/T$ (dBm)

- Signaux parasites :
  . Autoconvolution du signal "i" : $20 \log P_i - F - 20 \log \theta/T - (R + A/2 + 10 \log \beta\theta)$ (dBm)
  . Convolution double trajet : $10 \log(P_1 P_2) - F - 20 \log \theta/T - (2R + A)$ (dBm)
  . Bruit thermique : $-114 + 10 \log B$ (MHz)   (dBm)

F : facteur de bilinéarité (-72 dBm typ.)

R : coefficient de réflexion (10-20 dB si transducteur classique)
                                  > 30 dB si transducteur unidirectionnel)

A : perte de propagation acoustique (3-5 dB)

1.2.4. Exemples d'applications.

Pour démontrer les qualités du dispositif nous avons fait des expériences de compression d'impulsion c'est-à-dire que nous avons fait le produit de convolution de signaux linéairement modulés en fréquence avec leur réplique inversée dans le temps. La Fig. 7a correspond au cas de l'autocorrélation d'une rampe d'excursion de fréquence 24 MHz et de durée 12 μs (Produit BT = 288). La largeur à 4 dB est de l'ordre de 21 ns ce qui est voisin de la valeur théorique 20,8 ns. On note aussi l'absence d'échos parasites lointains.

Au moyen d'un doubleur de fréquence on atteint un BT théorique de 576. Les résultats sont montrés sur la Fig. 7b. La largeur à 4 dB est ici d'environ 11 ns en accord avec la valeur théorique 10,5 ns. Nous voyons que l'impulsion s'est légèrement dégradée par rapport à la précédente ce qui est dû à des défauts de phase sur les bords de la bande passante et aussi à certains défauts des signaux FM d'entrée aussi représentés sur la Fig. 7b. Mais la qualité du signal est très bonne.

Le convoluteur piézoélectrique permet donc d'effectuer la convolution ou la corrélation de signaux à très grande vitesse et sous un encombrement réduit. Les principales applications sont la détection, l'identification de signaux et la mesure de temps d'arrivée. Nous ne les détaillerons pas ici car elles font l'objet d'une autre présentation [12].

Le fonctionnement en corrélateur nécessite l'inversion dans le temps d'un des signaux, opération qui peut être effectuée au moyen de mémoires analogiques ou numériques [12]. En outre, ce composant peut être utilisé de façon asynchrone au prix d'une perte de capacité de traitement dans un rapport 2 ; mais il reste mieux adapté à une opération sychrone.

Pour s'affranchir de ces contraintes, un nouveau composant, le corrélateur à mémoire, est à l'étude.

2. - LE CORRELATEUR A MEMOIRE.

2.1. Principe.

La géométrie de ce composant est semblable à celle des convoluteurs à semiconducteur. Le dispositif se compose d'un milieu piézoélectrique couplé à un milieu semiconducteur qui est formé d'un réseau de diodes intégrées. Celles-ci sont du type Schottky ou pn et leur périodicité doit être inférieure à une demi longueur acoustique pour assurer un bon couplage. Le schéma d'un dispositif à diodes pn est représenté sur la Fig. 8. Le réseau de diodes sert d'une part d'élément non linéaire comme dans le cas du convoluteur simple, et d'autre part d'élément de mémoire où l'information est stockée sous forme de la charge de la capacité des diodes.

Pour décrire le mode de fonctionnement du dispositif, nous faisons référence à la Fig. 9. Soit $G(t) \cos \omega t$ le signal de pulsation ω que l'on veut inscrire dans le dispositif, nous supposons sa durée au plus égale au retard acoustique sous le réseau de diodes. Ce signal est appliqué sur le transducteur T_1 de manière à générer une onde acoustique. Lorsque cette onde est toute entière en regard du semiconducteur on applique une impulsion $W(t)$ très brève et de grande amplitude sur l'électrode de plaque E de manière à polariser fortement en direct les diodes. Celles-ci se chargent alors quasi-instantanément et la distribution de charges ainsi stockée $q(z)$ contient une composante spatialement uniforme et une composante de périodicité égale à celle de l'onde acoustique et d'amplitude proportionnelle au signal acoustique. A condition que la durée τ_F de l'impulsion dite d'écriture soit inférieure à une demi-période de RF ($\tau_F < \pi/\omega$), on peut écrire :

$$q(z) = q_o + a\, G(-z/v) \cos \omega z/v$$

Une méthode similaire d'écriture consiste à appliquer sur l'électrode de plaque E un signal à la pulsation ω, $W(t) \cos \omega t$ qui créera donc un battement stationnaire avec le signal acoustique qui sera inscrit à son tour dans le réseau de diodes. Le phénomène mis en jeu est maintenant une interaction cohérente aussi l'information sera correctement stockée à condition que la durée τ_D du signal W soit inférieure à l'inverse de la

largeur de bande de $G(t)$. Les signaux utilisés dans les dispositifs acoustiques ayant une largeur de bande relative au plus égale à 40 %, la durée de l'impulsion d'écriture est donc au moins 5 fois plus grande que si l'on emploie la technique précédente.

L'écriture une fois terminée, les diodes se retrouvent polarisées en inverse sous l'effet de la charge stockée ; elles ne se déchargent donc que très lentement au rythme de leur courant inverse et l'information est conservée en mémoire pendant un temps qui peut atteindre, à température ambiante, quelques dizaines de millisecondes ou quelques secondes s'il s'agit de diodes Schottky ou de diodes pn respectivement. Remarquons que dans cette opération l'amplitude et la phase du signal acoustique ont été stockées.

Si l'on applique un signal de même fréquence $F(t) \cos \omega t$ sur le transducteur T_1, on peut relire l'information stockée par mélange non linéaire dans le semiconducteur. Le signal résultant est détecté sur la plaque à la pulsation ω. Il est de la forme $H(t) \cos \omega t$ avec :

$$H(t) \propto \int F(t - z/v) \, G(-z/v) \, dz$$

ou encore après le changement de variable $\tau = t - z/v$

$$H(t) \propto \int F(\tau) \, G(\tau - t) \, d\tau$$

c'est-à-dire le produit de corrélation de F et G.

Le système est ainsi un corrélateur à mémoire ou encore un filtre linéaire dont la réponse impulsionnelle est programmable.

2.2. Etat actuel des recherches.

Du fait du grand intérêt que suscite ce type de composant, les études se poursuivent dans plusieurs laboratoires en France, aux Etats-Unis et en Norvège [13 - 17]. Les travaux portent principalement sur l'étude du détail du fonctionnement et sur l'étude des divers types et géométries de diodes.

Les trois paramètres à examiner en priorité sont le temps d'écriture, le temps de mémoire et le rendement énergétique.

Pour que le système soit utilisé efficacement, il faut qu'au bout de la durée d'inscription, les diodes soient presque totalement chargées. La durée du signal d'écriture définit aussi, comme nous l'avons vu, la valeur maximum soit de la largeur de bande du signal à inscrire $\Delta f < 1/\tau_p$, soit celle de la fréquence de ce signal, $f < 1/\tau_F$, et donc indirectement sa largeur de bande $\Delta f < 0,4/\tau_F$, selon la méthode d'écriture employée. On montre que la rapidité d'écriture est d'autant plus grande que l'amplitude du signal d'écriture est forte. Par exemple, si l'on utilise la méthode d'écriture par impulsion brève, celle-ci doit avoir une centaine de volts pour que l'inscription soit efficace et pour que le temps d'écriture atteigne 1 ns [13, 18] . Si la méthode d'inscription paramétrique est utilisée et si la puissance de l'impulsion C.W. est de l'ordre de 1 W, le temps d'inscription est de l'ordre de 0,5 µs avec des diodes pn, mais il est de l'ordre de 10 ns avec des diodes Schottky. Cette méthode autorise donc l'utilisation d'impulsions C.W. d'amplitude modérée tout en permettant de traiter des signaux de largeur de bande d'au moins 100 MHz.

Si l'on mesure la variation de l'amplitude du signal de corrélation H avec le temps qui sépare l'instant d'inscription de celui de la lecture, on peut définir un temps de mémoire τ_{st} au bout duquel H a diminué de 1 dB. Pour les dispositifs à diodes pn, τ_{st} est de l'ordre de 1 s avec une valeur extrême de 10 s avec des diodes de type vidicon [14] . Pour les dispositifs à diodes Schottky de périodicité 12,5 µm, nous avons obtenu des valeurs de τ_{st} d'environ 80 ms [15]. (Ceci correspond aux meilleurs résultats publiés à ce jour). Tous ces résultats ont été mesurés à température ambiante.

Comme pour les convoluteurs, on peut définir un facteur de bilinéarité F_M :

$$F_M \text{ (dBm)} = 10 \log \left\{ P_{correlat.} \text{ (mW) } / \, P_F \text{ (mW) } P_G \text{ (mW)} \right\}$$

qui dépend cependant du signal d'écriture W. Il existe néanmoins en général un optimum et F_M varie entre -70 et -80 dBm selon les dispositifs. On peut aussi définir la perte d'insertion du système utilisé en filtre à corrélation, c'est-à-dire le rapport P_{corr}/P_F pour $P_G = 1$ W. La perte est alors d'environ 50 dB.

Ceci signifie que le signal maximum de sortie est de l'ordre de -30 dBm si l'on suppose des puissances électriques d'entrée de 1 W. Le rapport signal à bruit thermique de sortie avoisine donc les 60 dB. Il faut aussi compter sur la présence de signaux parasites dus au signal de lecture F et qui sont à l'heure actuelle seulement 25 dB au-dessous du pic de corrélation.

Les principales caractéristiques des dispositifs existants sont rassemblées dans le Tableau IV ainsi que les limitations envisagées. Rappelons que tous ces dispositifs emploient la technologie dite "à film d'air" (voir Fig. 3a). Une version monolithique du type de celle représentée sur la Fig. 3c est également à l'étude et les premiers résultats semblent très encourageants [19] .

2.3. Applications au traitement par corrélation.

L'application première de ce genre de dispositif est la possibilité de corréler deux signaux F et G de durée limitée mais asynchrones. Pour démontrer ce point nous avons utilisé le corrélateur à diodes pn dont les caractéristiques ont été rassemblées dans la première colonne du Tableau IV. Nous considérons deux signaux identiques de durée 6 µs, mais décalés dans le temps l'un par rapport à l'autre de T = 10 ms. Le signal G est tout d'abord inscrit et stocké dans le dispositif. Le système est alors un filtre adapté au signal G ou F ; il peut donc filtrer le signal F quand il se présente. Le résultat est le produit d'autocorrélation de F.

- TABLE IV -

Caractéristiques des corrélateurs à mémoire

Type de diode	Démontré			Limites	
	pn	Schottky		Schottky	pn
Périodicité diode	12,5 µm	12,5 µm	5,8 µm	4 µm	
Fréquence centrale (MHz)	120	120	145	300	
Bande passante (MHz)	20	20	30	100	
Durée signaux (µs)	6	6	10	10-20	
Perte en filtre à corrélation (dB)	-50 (*)	-50 (*)	-50 (•)	-	
Signal/parasite sortie (dB)	25	25	25	40	
Signal/Bruit sortie (dB)	60	60	60	60	
Temps mémoire	1-10 s	80 ms	10 ms	0,1 s	1 s
Référence	[14]	[15]	[11]		

(*) Mesuré pour P_W = 1 W

(•) Mesuré pour W = 300 V en 5 ns

La Fig. 10a représente le cas où le signal F est une rampe pure, de durée 6 µs et d'excursion en fréquence 12 MHz. La trace supérieure correspond au signal FM, la deuxième et la quatrième trace à la sortie autocorrélation. La présence de signaux parasites perturbe les lobes secondaires. Ces signaux sont visibles sur la troisième trace, qui correspond au signal de sortie en l'absence de tout signal stocké. La Fig. 10b montre les résultats similaires obtenus lorsque le signal FM initial a subi un filtrage de bande. La pondération ainsi introduite élargit le lobe principal et abaisse le niveau des secondaires.

Ce nouveau type de composant est donc le filtre idéal au sens de la théorie de l'information. Il suffit de prélever le signal au niveau de l'antenne et de l'inscrire en mémoire dans le corrélateur où il sert alors de référence. Le dispositif se comporte alors en filtre programmé, adapté au signal réellement transmis. Si ce signal est variable dans le temps (radar agile), le récepteur restera correctement adapté à condition que le signal émis soit mis en mémoire suffisamment souvent.

Outre cette application, le type de composant doit trouver son utilisation dans le stockage et l'analyse rapide de signaux mémorisés (filtrage, identification, analyse spectrale...). Il combine les fonctions mémoire et corrélation, et il est ainsi très souple d'emploi. Des exemples d'application seront décrits dans un autre article présenté dans le cadre de cette réunion [12].

- - - - - - - -

Le convoluteur piézoélectrique à compresseur de faisceau fournit une solution simple au problème de la convolution rapide de signaux analogiques de produit B.T de plusieurs centaines et de largeur de bande de plusieurs dizaines de Mégahertz. Il apparaît donc bien adapté au traitement de signaux radar. La réalisation est simple car sa géométrie planar ne nécessite qu'un seul niveau de masquage.

Un nouveau composant acoustique combine les fonctions mémoire et corrélation, rendant alors possible une opération asynchrone du type filtre adapté programmable. Un tel dispositif peut donc grandement simplifier la réalisation de systèmes de traitement du signal. Aussi son étude se poursuit-elle activement.

-:-:-:-:-:-:-:-:-

- REFERENCES -

[1] C.F. QUATE and S.B. THOMPSON "Convolution and Correlation in real time with Non Linear Acoustics" Appl. Phys. Letts. 16, 494-496 (1970)

[2] L.O. SVAASAND "Interaction between Elastic Surface Waves in Piezoelectric Materials", Appl. Phys. letts. 15, 300-302 (1969)

[3] M.V. LUUKKALA and G.S. KINO "Convolution and Time Inversion using Parametric Interactions of Acoustic Surface Waves", Appl. Phys. Letts. 18, 393-394 (1971)

[4] Ph. DEFRANOULD and C. MAERFELD " A SAW Planar Piezoelectric Convolver", Proc. IEEE, 64, 748-751 (1976)

[5] W.C. WANG and P. DAS "Surface Wave Convolvers via Space Charge Non linearity" Proc. 1972 IEEE Ultras. Symp. 316-321 (1972)

[6] O.W. OTTO and N.J. MOLL "A Lithium Niobate-Silicon Surface Wave Convolver" Electron. Lett. 8, 600-602 (1972)

[7] G.S. KINO, W.R. SHREVE and H.R. GAUTIER "Parametric Interactions of Rayleigh Waves" Proc. 1972 IEEE Ultras. Symp., 285-287 (1972)

[8] S.A. REIBLE, J.H. CAFARELLA, R.W. RALSTON, E. STERN "Convolver for DPSK Demodulation of Spread Spectrum Signals" Proc. 1976 IEEE Ultrasonic Symp. 451-455 (1976)

[9] T.W. GRUDKOWSKI and C.F. QUATE "Acoustic Readout of Charge Storage on GaAs" Appl. Phys. Letts. 25, 99-101 (1974)

[10] B.R. KHURI-YAKUB and G.S. KINO "A Monolithic Zinc-Oxide on Silicon Convolver" Appl. Phys. Letts. 25, 4, 188-190 (1974)

[11] T.M. REEDER, M. GILDEN "Convolution and Correlation by Nonlinear Interaction in a Diode-Coupled Tapped Delay Line" Appl. Phys. Lett. 22, 8-10 (1973)

[12] P. ANTHOUARD, T. BEAUVAIS "Applications du convoluteur pizéoélectrique au traitement du signal radar" Paper 5.4 presented at the 34th Technical Meeting of the Avionics Panel of AGARD, Ottawa 10-14 Oct. 1977

[13] K.A. INGEBRIGTSEN "The Schottky Diode Acoustoelectric Memory and Correlator - A Novel Programmable Signal Processor", Proc. IEEE 64, 5, 764 (1976).

[14] Ph. DEFRANOULD, H. GAUTIER, C. MAERFELD, P. TOURNOIS "Pn Diode Memory Correlator" Proc. 1976 IEEE Ultrasonics Symp., 336-347 (1976)

[15] H. GAUTIER, Rapport interne Thomson-C.S.F. DASM (1977)

[16] P.G. BORDEN, G.S. KINO "Correlation with the Storage Convolver" Appl. Phys. Lett. 29, 9, 527 (1976)

[17] K.A. INGEBRIGTSEN, A. RØNNEKLEIV "Experiments with Acoustoelectric Memory Correlators" Paper 3.6 presented at the 34th Technical Meeting of the Avionics Panel of AGARD, Ottawa, 10-14 oct. 1977.

[18] P. BORDEN, G.S. KINO "The Charging Process in the Acoustic Surface Wave p.n Diode Storage correlator" Ginzston Laboratory Report, Stanford University (USA) (July 1977)

[19] H.C. TUAN and G.S. KINO "A Monolithic Zinc-Oxide-on-Silicon p.n Diode Storage Correlator" Ginzston Laboratory Report, Stanford University (July 1977).

(a) $\quad$ F(t) $\longrightarrow$ [Filtre G(t)] $\longrightarrow$ $H(t) = \int F(\tau) G(t-\tau) d\tau$

(b) $\quad$ F(t) $\longrightarrow$ [Convo- luteur acoustique] $\longrightarrow$ $H(t) = \int F(\tau) G(2t-\tau) d\tau$

$\qquad\qquad\qquad$ G(t)

(c) $\quad$ F(t) $\longrightarrow$ [Dispositif à mémoire G(t)] $\longrightarrow$ $H(t) = \int F(\tau) G(-t+\tau) d\tau$

$\qquad\qquad$ mémorisation
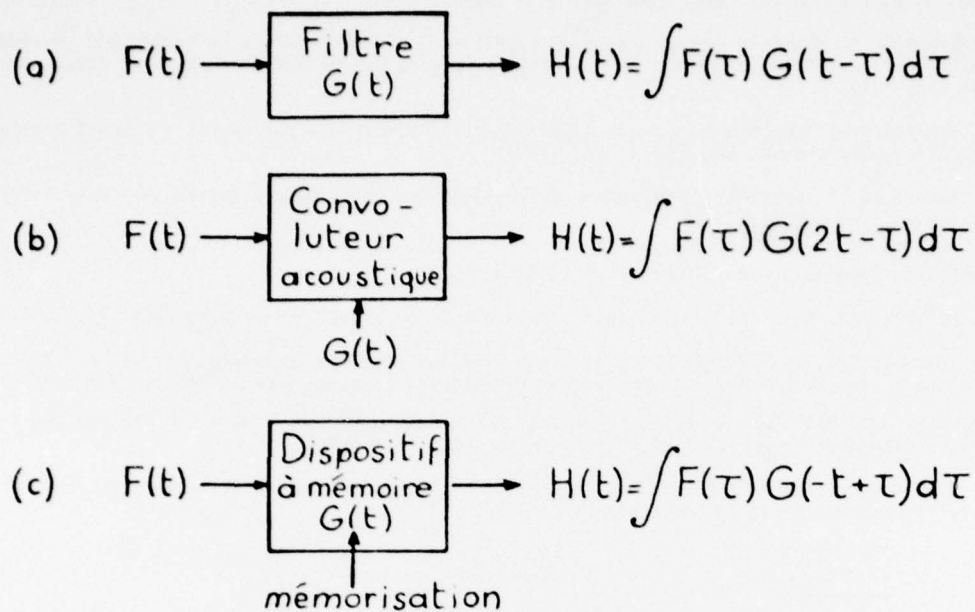
Fig. 1 - Les fonctions filtre linéaire, convolution et corrélation avec mémoire.

Convolution
F*G à 2ω

Paire d'électrodes

L

Signal
F(t)e$^{j\omega t}$

Signal
G(t)e$^{j\omega t}$

z

Milieu piézoélectrique
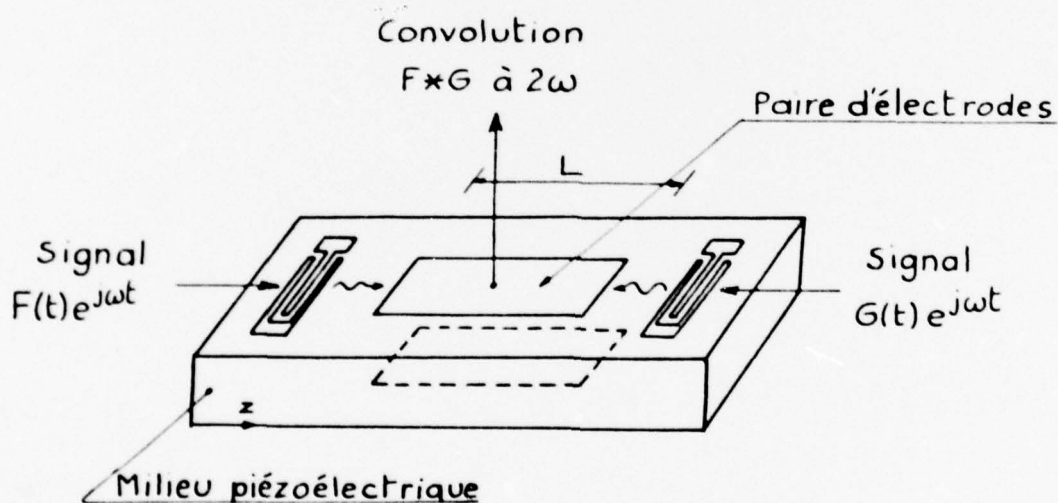
Fig. 2 - Principe de fonctionnement du convoluteur piézoélectrique à onde acoustique de surface.
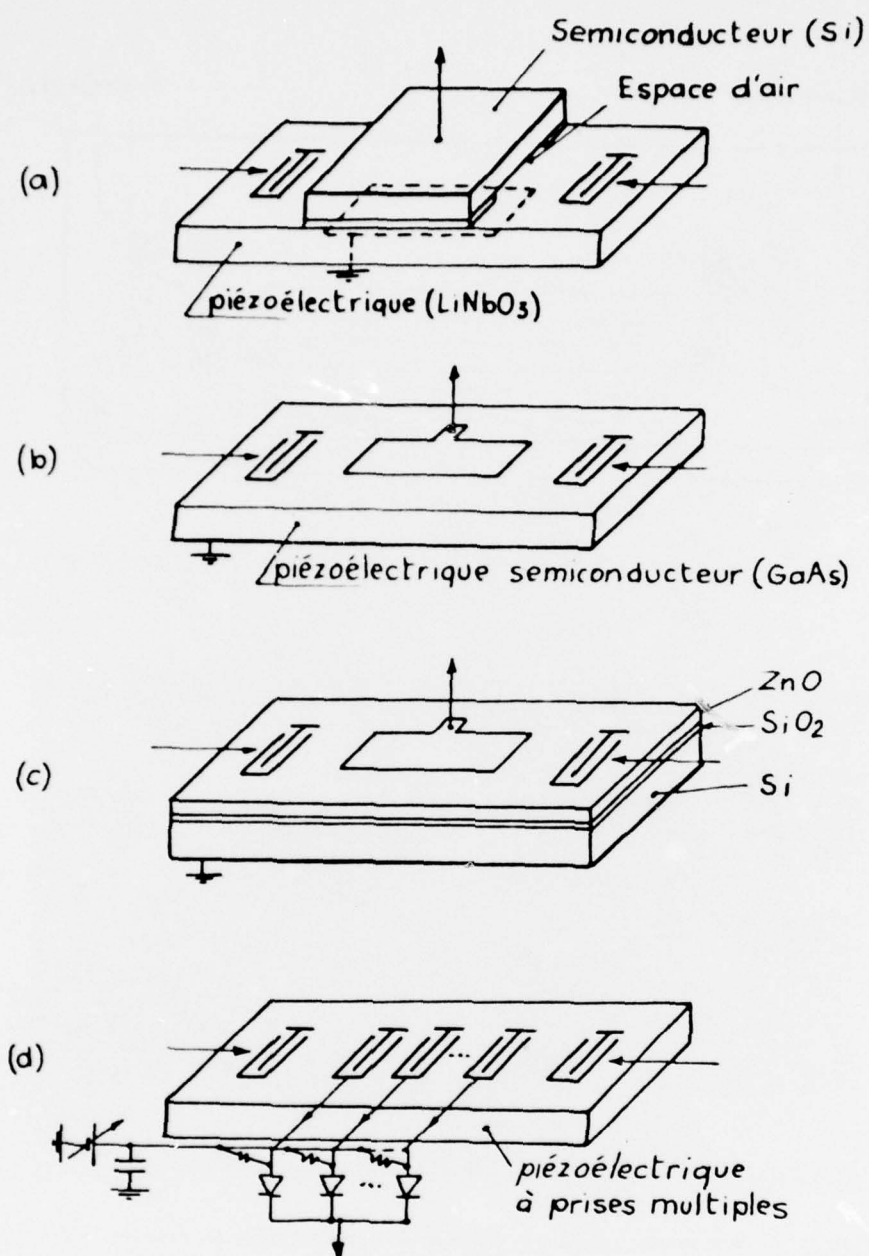
Fig. 3 - Principaux types de convoluteurs à semiconducteur.
a) Convoluteur à milieux séparés [5, 6, 7, 8]
b) Convoluteur à piézoélectrique semiconducteur [9]
c) Convoluteur à couche mince piézoélectrique [10]
d) Convoluteur hybride à réseau extérieur de diodes [11]
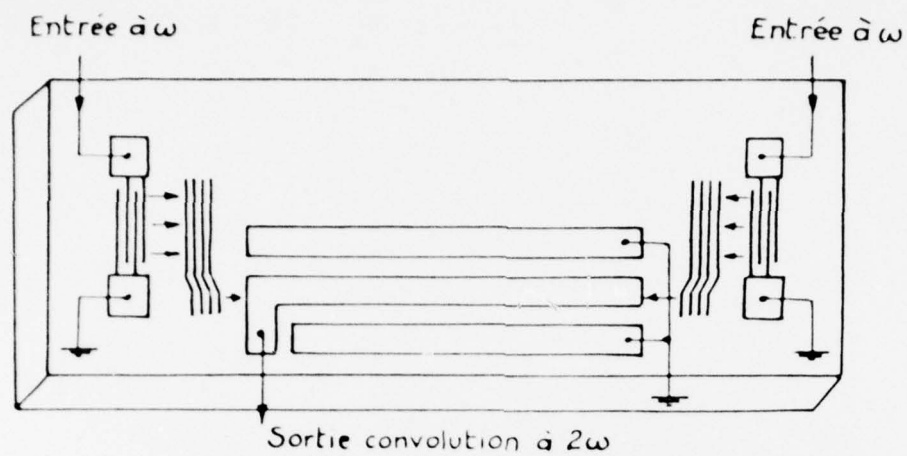
Fig. 4 - Schéma du convoluteur piézoélectrique à compresseur de faisceau
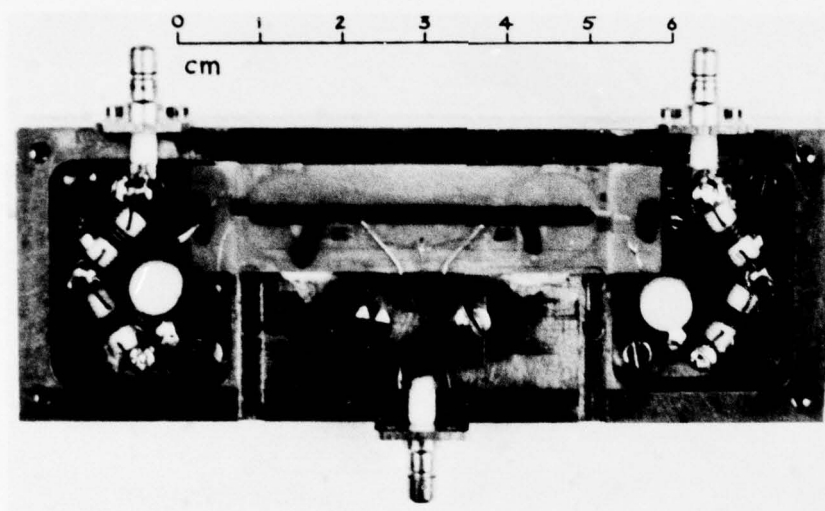
Fig. 5 - Photographie d'une réalisation d'un convoluteur du type de la Fig. 4
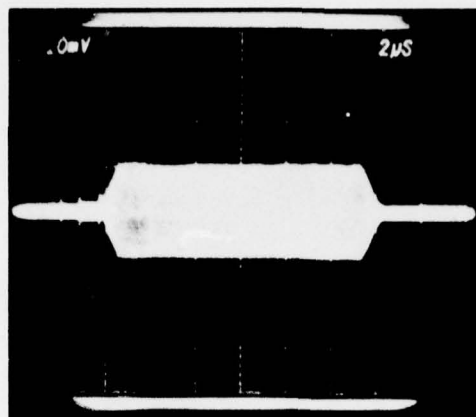
Fig. 6 — Signal de convolution d'une impulsion large avec une étroite, démontrant l'uniformité de la réponse.

(a)

(b)
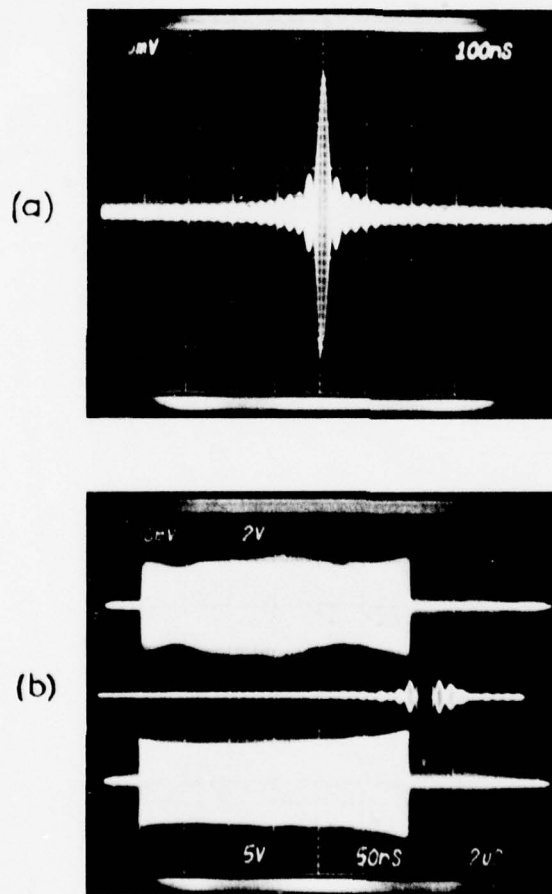
Fig. 7 — a) Autocorrélation d'une rampe FM d'excursion de fréquence 24 MHz et de durée 12 µs (Largeur à 4 dB = 21 ns)

b) Autocorrélation d'une rampe FM d'excursion de fréquence 48 MHz et de durée 12 µs. Les signaux d'entrée sont représentés sur les traces supérieure et inférieure. Le signal de convolution est représenté sur la trace intermédiaire (largeur à 4 dB = 11 ns).

n Si

p⁺ p⁺

SiO_2

12.5 μm

contacts

E

T_1

Y

Z

$LiNbO_3$
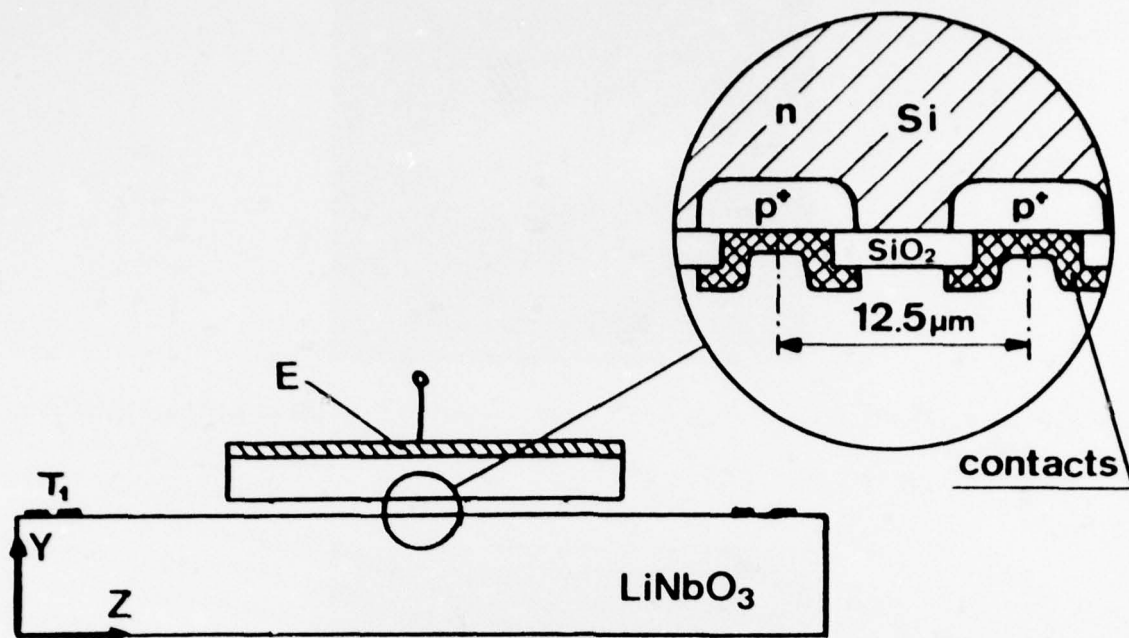
Fig. 8 - Schéma d'un corrélateur à réseau de diodes pn de type vidicon.

$w(t)$ τ_F ①

$w(t)e^{j\omega t}$ τ_P ②

$q(z) = q_0 + a\, G\left(-\frac{z}{v}\right)\cos\frac{\omega z}{v}$

E

(a)

$G(t)\,e^{j\omega t}$

T_1

z

$e^{j\omega t}\int F(\tau)G(-t+\tau)d\tau$

E

(b)

$F(t)\,e^{j\omega t}$

T_1
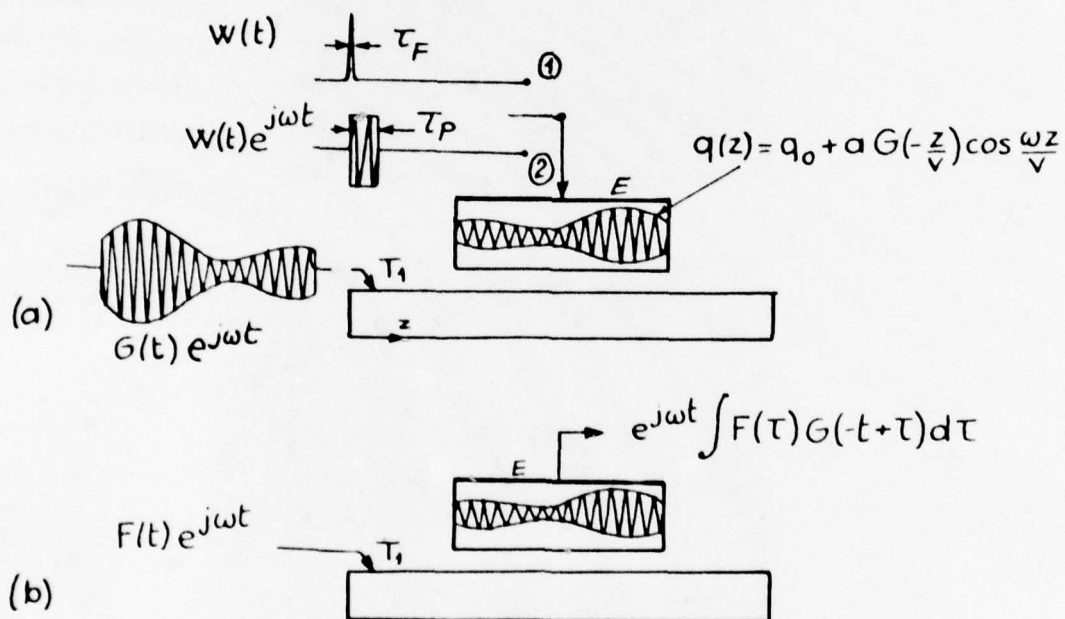
Fig. 9 - Principe d'utilisation du corrélateur à mémoire
   a) Inscription du signal à mémoriser G
   b) Lecture par corrélation asynchrone
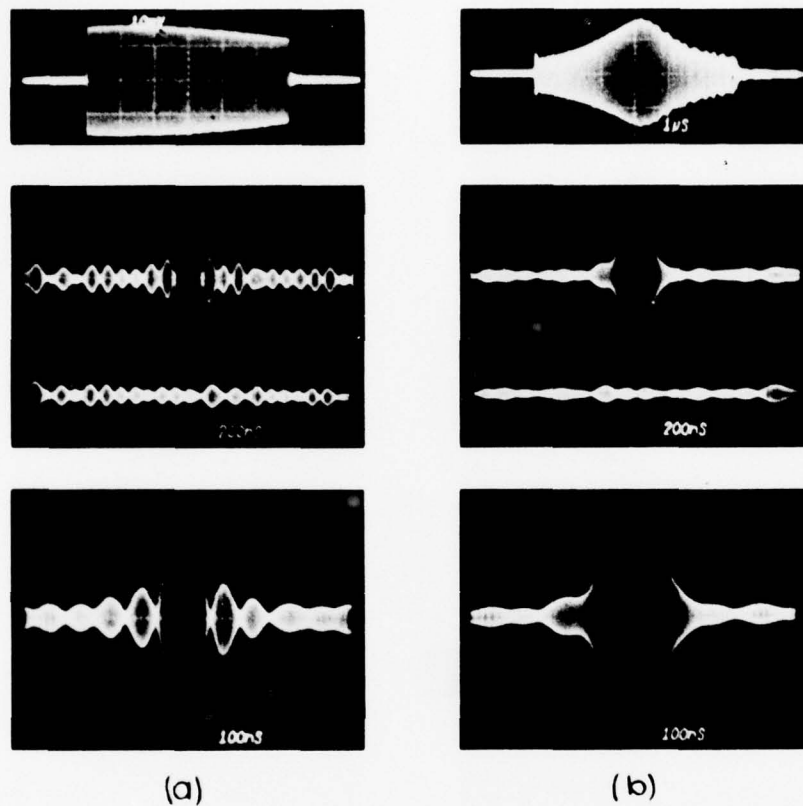
Fig. 10 - a) Corrélation d'une rampe de fréquence de durée 6 µs, bande de fréquence 12 MHz avec sa réplique stockée 10 ms plus tôt.

b) Corrélation du même signal ayant au préalable subi un filtrage de bande (Trace supérieure : Signal FM utilisé - Traces n° 2 et n° 4 : sortie autocorrélation Trace n° 3 : sortie en l'absence de signal mémorisé).

DISCUSSION

Jackson

I am not quite sure if I understand the full benefit of beam compression. Do you have any degradation in gain bandwidth product?

Réponse d'Auteur

La compression de faisceau ne fait qu'augmenter la densité de puissance. Pourquoi nous pouvons l'augmenter? C'est parce que nous ne sommes pas limités par la saturation de l'effet non linéaire dans un substrat seulement piézoélectrique. Ainsi en augmentant la densité de puissance dans une large bande, ce qui se fait grâce au compresseur de faisceau, nous n'avons pas de dégradation, mais simplement un gain en efficacité.

Question non enregistré

Réponse d'Auteur

La capacité de sortie est réglée par le système triplaque que vous avez vu et dans un tel système lorsque les électrodes sont dispersés de part et d'autre de l'électrode centrale on peut montrer qu'à rapport de géométrie identique la capacité de sortie ne dépend pas de la largeur.

LSI VIDEO COMPRESSION AND COMPUTATIONAL MODULES
UTILIZING DIGITAL CHARGE COUPLED DEVICES

R. A. ALLEN, D. J. SPENCER AND C. S. MILLER
TRW DEFENSE AND SPACE SYSTEMS GROUP
REDONDO BEACH, CALIFORNIA

SUMMARY

During the last two years, a family of digital devices have been developed that utilize the CCD technology. The logic family consists of AND, OR, and exclusive-OR gates, INVERT, HALF and FULL-ADDERS, shift-registers and interface circuits. The techniques used in the design of LSI arithmetic arrays are described in this paper and is referred to as Digital Charge Coupled Logic (DCCL). The DCCL technique maintains the high density implementations and low speed-power product of CCD's and adds the advantages of high noise immunity, freedom from device variation and absolute calculation accuracy inherent with digital operation.

This paper describes an LSI Video Compression Module concept utilizing the DCCL technology, that compresses video data by intraframe transform coding. The algorithm chosen for primary implementation is the two-dimensional Hadamard transform with block quantization.

This paper also describes the design of the DCCL arithmetic and control chips. Several applications are discussed. In particular is the use of these two types of DCCL LSI chips in the realization of the Itakura analysis - synthesis algorithm for a linear predictive coding (LPC) using partial correlation (PARCOR) techniques.

This work has been supported in large part by the United States Naval Research Laboratory, under contract number N00014-74-C-0068.

INTRODUCTION

With the exception of shift-register memories, almost all CCD's applications have been in analog implementation of signal processing systems. However, during the last three years a few papers (MILLER, C. S., 1975; ZIMMERMAN, T. A. et al, 1977) have described the use of CCD's to perform pipelined arithmetic. The details of the basic CCD technology, high gate density, simple processing steps, and very low power dissipations are well known to the participants at this symposium. However, a comparison between the difference in the way in which an analog approach represents a signal, and the way in which a digital approach represents a signal, has some merit.

In the analog approach, one signal processing sample is used for each signal sample and a significance is attached to the signal amplitude, but in the digital approach a n-bit word is used to represent each signal sample and the amplitude significance is contained within the n-bit word. Any loss of carriers from the charge packet due to transfer inefficiency or accumulation of thermal carriers by the charge packet has a significantly different effect in the analog and digital domain. In the analog approach, changes to the size of the charge packet representation changes the modulation transfer function which results in a degradation of both the amplitude accuracy and the phase relationship as described(JOYCE, W. B. et al, 1971). In the digital approach we assign two threshold levels to the charge packet (binary one and binary zero). Providing the loss of carriers does not cause the binary "1" packet of charges to fall below the high threshold, or the increase in the binary "0" does not exceed the low threshold then there will be no degradation of the processed data.

DIGITAL CHARGE COUPLED LOGIC (DCCL)

The Digital Charge Coupled Logic (DCCL) implementation of the OR, AND, ADD functions have been reported earlier (ALLEN, R. A. et al, 1976; ZIMMERMAN, T. A. et al, 1977). However, since these basic logic gates are an integral part of any higher level function,we will describe them briefly again in this paper. We shall consider that a binary "1" consists of a charge packet that contains a predetermined number of carriers and a binary "0" consists of an empty packet (no carriers).

In the overlapping gate DCCL structure shown in Figure 1, three separate input CCD channels A, B, G, converge under one common storage area D. The size of the D charge bucket is designed so that it will be completely filled if a binary "1" packet is transferred into it from either of the three input channels.

There are two output ports from the D storage area, controlled by gates b_1 and s_1. When gate s_1 is switched to a transfer mode, the charge packet stored under D transfers out producing the OR function in Boolean symbols $S_1 = A + B + G$.

If any two of the input channels transfer a binary "1" packet under the D storage area, it will fill completely with half of the total input charge so that the remaining charge overflows across the potential barrier b_1 to fill the M storage area. When gate X switches to a positive value the charge packet under M transfers out as in Boolean symbols $X = AB + AG + BG$.

For the third condition when all three input channels transfer binary "1" charge packets under the D storage area, the total charges fill D, overflow across b_1 to fill M and again overflow across the b_2 barrier and then also fill the I storage area. Subsequentially when the S_2 gate is switched to a positive level, the charge packet under the I storage area transfers out as the AND function, $S_2 = ABG$. Thus with one simple DCCL gate we have simultaneously produced the AND, OR and a combinational logic function.

Since DCCL is a dynamic system, the D, M, and I storage areas must be completely emptied of charge between input changes. If any of the three specific functions are not required for further signal processing the charge packet at that exit port must terminate at a sink diode.

Logic functions such as an exclusive-OR require the complement of a function. Logic inversion is implemented in DCCL by the use of a floating-gate amplifier. Earlier papers (BARBE, D. F, 1975; TOMPSETT, M. F. 1972) have described a floating-gate amplifier in which the present potential level is induced from a bias electrode overlapping the floating-gate. During development of DCCL circuits we have found it preferrable to connect the voltage directly to the floating-gate through a FET used as a switch. This also enables us to use the source of the FET as a floating- diffusion and also to sink the charge on the floating-diffusion at the end of a clocking sequence.

The layout of a DCCL exclusive-OR gate is shown in Figure 2. As described above, if either of the two input channels transfers a binary "1" into the D storage area, it will transfer across t and s_1 to produce the function $S1 = A\bar{B} + \bar{A}B$. However, when both input channels transfer binary "1" packets into the D storage area so that it overflows across b_1 to fill M, the resulting change in surface potential under M induces a potential change on the floating-gate or floating diffusion. The potential change is transferred to the s_1 end of the floating-gate causing it to switch from a transfer level to a charge barrier level. Now when t is switched to a positive voltage, the charge packet transferring from D will be retained under t by the s_1 acting as a barrier. During the next clock phase when c is switched to a positive voltage, the packet of charges held in t will transfer out under c to produce the function $C = AB$ and no charges will transfer out of s_1.

The DCCL exclusive-OR is identical to a half-adder in which the S1 output is the SUM and the C output is the CARRY to the next level. A full-adder can be implemented by increasing the input channels to three, and by adding the b_1 and I areas as shown in Figure 1. The S1 and S2 are connected together in a DCCL OR gate to form the SUM output. A photograph of a DCCL (16 + 16) adder array utilizing 16 full-adders, carry-synchronizing and deskewing shift registers is shown in Figure 3.

Two questions are often raised about DCCL. Firstly, can long logic chains be implemented even through the signal level degrades due to transfer losses? Secondly, can each function fan-out to more than one other function? The answer to both of these questions centers on the use of a digital refresh cell.

It is a simple matter to modify a half-adder DCCL cell so that it will perform the digital refresh function. All that is required is that a diode and fill and spill gate be arranged to insert a binary "1" into the D storage cell on each clock phase, synchronous with the signal charge packet in the other channel.

If the B input channel is designated the continuous "1" channel and channel A the input signal channel, then the SUM and CARRY outputs become $S = \bar{A}$ and $C = A$. Since the C = A charge packet is a binary "1" only when D overflows, it must be a completely full charge packet no matter what the quantity of charges in the original A input; thus A is refreshed. Two rows of such digital refresh cells are shown in Figure 3. This modified cell can also be used to generate the complement of the input signal.

In order to provide a fan-out capability, it is necessary to multiply the input charge packet. One approach is to multiply the fill and spill input areas of the refresh cell, by the fan-out required. The C output from the t gate is then divided into the number of channels required for the fan-out.

In the two signal filtering applications described in this paper, it is necessary to frequency divide and also to store a signal in latch registers. The digital refresh cell can perform both these functions.

By connecting the S1 signal back to an OR-gate at the A input, charge packets will be generated at the C and S1 output ports that are one half the clock frequency. If a 1-bit shift-register delay is inserted in the feedback between S1 and A, the charge packet train generated will be one quarter the clock frequency. The clock rate can thus be divided down by any even number by inserting the appropriate delay in the feedback loop.

To implement a latch circuit it is necessary to feedback the C output from a refresh cell to an OR gate at the A input. If the initial SET input to the latch is a binary "1", the latch will automatically refresh the charge packet each clock period but if the initial SET input is a binary "0", the latch will sink any thermal charges accumulated between clock periods.

LSI FOR VIDEO DATA COMPRESSION

Data communications has become an important complement to military avionics in many command and control applications. Most data transmission is today accomplished over digital communication systems taking advantage of their greater efficiency and superior error control. Television, however, has continued to utilize analog signalling techniques. This arises from the high data rates (typically 60 Mbps) and concomitant wide bandwidths required by conventional Pulse Code Modulation (PCM) of video. Typical parameters required for good fidelity include sampling rates of 10 MHz or greater with 6 to 8 bits per sample. Nevertheless, the switching, storage and processing advantages of digital data have led many commercial and military applications to favor digital television transmission. Fortunately, developments in digital television coding techniques have shown that significant compression of the high data rates can be achieved through exploitation of spatial, temporal and spectral (color) redundancy in color video data.

An extensive video data compression literature exists (PRATT, W. K., 1973) covering a broad range of techniques from simple sampling schemes to sophisticated statistical algorithms. Transform image coding offers one promising approach currently in use in prototype systems (HELLER, J. A. 1974; KNANER, S. C. 1975). Transform coding algorithms perform a one-or two-dimensional unitary transformation on the input data followed by some form of quantization of the resulting coefficients. At the receiver the inverse transform is performed to restrict the original data within some level of degratation. The Hadamard, Fourier, Haar, Slant and Cosine transforms all possess qualities desirable for data compression.

The chief disadvantages of transform coding are implementation complexity related to the large data block sizes needed for effective decorrelation and high sampling rates. With two (spatial) dimensional data blocks, which are favored for taking advantage of all spatial picture correlation, the computation rates become intense. For example, using the straightforward FFT, an 8 x 8 sample data block requires 3 "butterfly" operations per sample. Besides the complex multiply and add functions, special logic must be included to limit numerical noise effects. Such systems are practical only through use of LSI to reduce complexity. DCCL provides an excellent potential solution by its high functional density, low power and match to pipeline algorithms. An ideal device for significantly reducing computational complexity would perform one or more stages of a pipeline version transform algorithm. For general applicability, sufficient data range

would be provided to allow the same device to be used at any stage of either the forward or reverse transform for data block sizes as large as 16 x 16. As much of the clocking and control as possible should be included on chip.

As an example, consider the implementation of a Hadamard transform in DCCL pipeline technology. Figure 4 indicates two basic stages of a pipeline Hadamard transform, which is seen to involve only add, subtract and delays.

One crucial problem in generating the Hadamard transform of a signal is dynamic range increase. The Hadamard coefficients are generated by adding and subtracting outputs of previous stages which has the effect of doubling the dynamic range of the data for each successive stage. In a preliminary study to determine the feasibility of designing a dual function (16 x 1) Hadamard transform on a single chip, we showed that it could best be carried out by dividing each word into its separate bits and treating each bit as a repeatable Hadamard transform cell (HTC). The number of required HTC's then increases at each stage with the dynamic range.

The interconnection of four HTC's is shown in Figure 5. Each cell consists of two full-adders, four multiplexing AND gates and a feedback delay. The I/O Mux signals in a HTC A5 (first Hadamard stage, 5th significant bit) or HTC A6 require a single clock phase delay, while in the B5 and B6 stages, they require two clock phase delays. These delays are part of the Hadamard encoding technique and double at each stage as the data progresses across the chip.

The carry-bit from Stage A5 transfers vertically to Stage A6 after a one-bit delay through the full-adder and the input data to Stage A6 must also be delayed by one-bit to ensure they arrive synchronously with the carry-bit. This skewing of the data is inherent in any DCCL pipeline array to accommodate the finite charge transfer times.

DCCL COMPUTATIONAL MODULES

The Hadamard transform described in the previous section is implemented with DCCL adder and subtractor arrays. However, by using additional arrays of adders, the high density DCCL technology is capable of performing many pipeline arithmetic functions on the same chip.

These arithmetic capabilities can be utilized in filtering applications: correlation, convolution, and fast transforms (i.e., Fourier, Hadamard, Hilbert, etc.). The inherent structure to these functions allows them to be cast into flow-form which is ideal for DCCL pipeline computations. The flow form of the structure allows timing and data routing to replace much of the program memory and control logic found in general purpose processors.

A large number of different flowform signal processing functions can be formed using only two types of DCCL Large Scale Integrated (LSI) chips. One chip provides the arithmetic function (AU) while the other provides memory and control (MC).

The basic arithmetic functions to be realized are addition/subtraction, multiplication and scaling (multiplication by a power of two). The arithmetic accuracy required for the different applications, as we have seen, can vary widely; however, the more stringent applications such as Itakura voice processing, can be satisfied with 16-bit multiplication accuracy and addition/subtraction accuracy of 16-bits, in some cases a double precision add/subtract capability of 32-bits is required. A DCCL chip configuration with this capability is shown in Figure 6. To allow the sequence of arithmetic operations to be performed in different orders corresponding to different applications, multiplexers are placed at the input to the adder and multiplier. A multiplexer is also provided at the output so that results of the different operations can be selected.

Control inputs are accepted by the arithmetic chip to route the data through the desired elements so that a prescribed sequence of operations is performed. For example, the radix two form of the FFT kernel or "butterfly" can be accomplished with six add/subtract operations and four multiplications. Initially, the sum of the real parts of the complex inputs is computed and the result applied to the output for storage. Next, the sum of the imaginary parts is computed and applied to the output. The third and fourth steps consist of computing the difference between the real parts and the difference between the imaginary parts, respectively. These differences are then routed through the multiplier where they are multiplied by the sine and cosine "twiddle factors". Four passes through the multiplier are required to compute the four products. The final steps in the "butterfly" computation which complete the complex multiply are to sum one pair of products and to compute the difference of the other pair. These two results are then returned to storage. To perform a correlation computation, the control inputs to the arithmetic chip cause the output of the adder to be fed back to the input thereby operating as an accumulator. The sequences of the two input variables to be correlated are then applied to the multiplier and the product accumulated. Registers for latching the data words are not shown in Figure 6 because storage is implicit in the operation of DCCL.

The fundamental consideration in applying DCCL is the throughput delay. Best computational efficiency is obtained using pipeline techniques. Although in many of the important applications, the computations can be cast into a flow form suitable for pipelining, there is also a need for general purpose computer operations such as executing branching and jump instructions. Conditional instructions of this type require a comparison to be made before the next step in the program is determined. Due to the delay through the DCCL arithmetic logic, it is difficult to perform general purpose computing efficiently. With these characteristics of DCCL arithmetic in mind, it is important to tailor the control chip to match the characteristics of the AU.

The basic timing can be broken up into blocks of N clock intervals where N exactly matches the delay through the AU: typically this may be between 16 to 32 clock intervals. Control of the DCCL arithmetic unit can be divided into block and intra-block instructions. The intra-block instruction controls allow pipelined operations to be performed, like the FFT, where data words are added and then subtracted on successive clock pulses. By changing the block instructions, the intra-block instructions can be changed after each block of N clock pulses. For example, the AU can perform successive sums and differences on one block of N samples followed by a multiply and accumulate on the next block of N samples. Since the results of an arithmetic operation becomes available at the arithmetic unit output after N samples, branching, skip, and jump instruction can be performed at the block rate. Thus, the AU can either perform flow form types of computations at the high rate of general purpose computation at the lower (by a factor of 16 to 32) block rate.

The separation of the control of the AU into block and intra-block functions is inherent in the shift register type of architecture for the DCCL control chip as shown in Figure 7. The blocks of shift register memory may be either conventional data memory or shift register read only memory (ROM) where a metallization mask determines how many of the blocks are to be ROM as well as the contents of the ROM. Multiplexers at the input and output control the flow of data and program instructions to and from the AUs.

Each memory block in Figure 7 consists of shift registers and recirculation logic. Nominally, the block would be 16 or 32 samples long with 16 shift registers in parallel corresponding to a 16-bit/word parallel data format. The shift register block has a tap after N-1 samples so that data can be recirculated back to the input after either N or N-1 sample delays. This allows the memory blocks to be operated either as recirculating memory or as a delay line time compressor (DELTIC) so that as the data processes through the register, the oldest sample is replaced by the new input sample.

When the DCCL control chip is operated with the DCCL arithmetic unit, the two data outputs are selected by the multiplex gates and applied to the AU inputs, while the arithmetic unit output is fed back to the input MUX. The lower multiplex gate on the right in Figure 7 selects the control values from either the ROM shift registers or a data shift register; thus, the controller operation can proceed either according to a program stored in the ROM or be changed by input values. The input values can be from either an external source (e.g., an interrupt) or from values computed by the DCCL arithmetic unit.

When operated at the slower block rate, both the data and the next program control value can be time-interleaved so that both are computed in a single N clock interval block time. With a 4 MHz clock rate and N = 32, both operations can be accomplished in 8 μsec.

Secure voice processing is a potential application for the pair of DCCL chips. The problem, stated simply, is to digitize the speech signal so that the security feature can be implemented by performing numerical operations on the data. Unfortunately, digitizing the speech directly (e.g., using PCM) results in a greatly increased bit rate. In order to reduce the bit rate so that it can be transmitted over existing audio channels, it is necessary to perform the secure voice processing which, in essence, compresses the bandwidth by source coding.

One of several voice processing techniques is to employ a linear predictive coding (LPC) algorithm (ITAKURA, F. et al, 1972). The procedure for the analyzer part of the voice processor consists of passing a signal through ten identical stages (as shown in Figure 8). In each stage a Parcor (ITAKURA, F. et al., 1972) coefficient, K_i, $i = 1, 2, \cdots 10$, is determined and the ten coefficients obtained in this way, plus pitch and voicing information are the data which are sent to the receiver where the speech is reconstructed.

As indicated in Figure 8, the Parcor coefficients can be written

$$K_i = \frac{\rho_i - 1}{P_{i-1}}$$

where the ρ_i are correlation values and the P_i are the mean square forward prediction errors (or backward prediction errors). The ρ_i and P_i can be written

$$\rho_i = \sum_{i-1}^{N} A_{i-1} B_{i-1}$$

and

$$P_{i-1} = E\left[A_{i-1}^2\right]$$

where E denotes expectation. The ith stage forward and backward prediction errors, A_i and B_i are related to the previous stage values, and the Parcor coefficient, K_i by

$$A_i = A_{i-1} - K_i B_{i-1}$$

and

$$B_i = B_{i-1} \quad K_i A_{i-1}$$

In practice, the summation over N forming the correlation is replaced by lowpass filtering. The expected value of the prediction error is approximated by subtracting the product of the Parcor coefficient times the correlation from the value in the previous stage. A block diagram indicating the operations necessary to realize an Itakura analyzer stage is shown in Figure 8.

Due to the relatively long propagation delay through the DCCL arithmetic unit, and the desire to employ pipeline computation for best speed and efficiency, it is necessary to separate the computation into a number of steps. Each Itakura analyzer section can be computed with a sequence of eight passes through the arithmetic chip to perform the following operations:

● Multiply $A_{i-1} \cdot B_{i-1}$

● Perform lowpass filter computation

$$\cdot \quad p_{i-1} = 2^{-k}(A_{i-1} B_{i-1}) + (1-2^{-k})p_{i-2}$$

● Compute $\dfrac{1}{P_{i-1}}$

Two multiplies, an add, and a table lookup are required.

- Compute $K_k = \dfrac{\rho_{i-1}}{P_{i-1}} = \rho_{i-1}\left(\dfrac{1}{P_{i-1}}\right)$

- Update $P_i = P_{i-1} - K_i \rho_{i-1}$

- Update $A_i = A_{i-1} - K_i B_{i-1}$

- Update $B_i = B_{i-1} - K_i A_{i-1}$

In the above list of operations, each can be accomplished in a straighforward manner by applying the appropriate controls to the arithmetic chip. The divide operation can be implemented by a recursive algorithm, which requires a ROM table lookup to estimate the inverse value followed by two multiplications and an addition (HUTCHINS, S. E. et al., 1975). Typically, a conventional ROM would be used; however, a DCCL ROM could be used if the longer access time could be tolerated.

A primary difficulty in implementing the Itakura analyzer in an efficient pipeline fashion with DCCL is that the result of one pass through the AU must be available before the next can commence. This, of course, holds true for each of the ten stages in the Itakura analyzer. The solution to this problem is to employ interleaving (MILLER, C. S. et al., 1975). By accepting, say, a 10 sample delay, the first operation on the list for each stage can be computed during the N clock pulse interval corresponding to the propagation time through the AU. On the next pass through the AU the second operation on the list is performed for each of the ten interleaved stages. After the last operation of the first analyzer stage has been completed, the time position of the result is shifted one clock interval utilizing the variable length shift register (or DELTIC) in the control chip so that the time position corresponds to the interleaving for the second analyzer stage. This procedure is repeated to complete the ten-stage analyzer computation. Although this distrubiton in time of the computation of different analyzer sections and the interleaving of the operations would require a complicated indexing scheme with a general purpose computer, with the shift register structure of the DCCL control chip the required control becomes simple and straightforward.

CONCLUSIONS

The use of charge coupled devices to realize digital logic and arithmetic functions has led to a new logic family referred to as DCCL. Complex parallel arrays can be fabricated to realize arithmetic functions. Because of the high functionsl density achievable with DCCL, a powerful computational capability (i.e., a 16 x 16 bit multiplier plus a 32 + 32 bit adder) can be fabricated on a single LSI chip. An example is shown where an array of repeatable cells combining logic and adders are arranged to perform (16 x 1) forward or reverse Hadamard transforms. A second example discusses the use of an arithmetic chip combined with a memory and control chip to perform a variety of digital filtering and signal processing operations. The specific example given is the analyzer for the Itakura secure voice processing algorithm.

REFERENCES

ALLEN, R. A., HANDY, R. J. and SANDOR, J. E., 1976, "Charge Coupled Devices in Digital LSI", International Electron Devices Meeting.

BARBE, D. F., 1975, "Imaging Devices Using the Charge-Coupled Concept", Proceedings of the IEEE, Vol. 63.

HELLER, J. A., 1974, "A Real Time Hadamard Transform Video Compression System Using Frame-to-Frame Differencing", NTC 1974.

HUTCHINS, S. E., et al., 1975, "Voice Processor LSI/Organization Optimization Study", TRW Final Report Number DCA 100-75-C-0020

ITAKURA, F. and SAITO, S., 1972, "On the Optimum Quantization of Feature Parameters in PARCOR Speech Synthesizer", IEEE Speech Conference.

JOYCE, W. B., and BERTRAM, W. H., 1971, "Linearized Dispersion Relation and Green's Function for Discrete Charge Transfer Devices with Incomplete Transfer", Bell System Technical Journal.

KNANER, S. C., 1975, "Real Time Video Compression Algorithm for Hadamard Transform Processing", Proc. of Society of Photo-Optical Instrumentation Engineers.

MILLER, C. S., and ZIMMERMAN, T. A., 1975, "Applying the Concept of a Digital Charge Coupled Device Arithmetic Unit", Second International Conference on the Application of Charge Coupled Devices.

PRATT, W. K., 1973, "Bibliography on Digital Image Processing and Related Topics", University of Southern California USCEE Report 453.

TOMPSETT, M. F., 1972, "A Simple Charge Regenerator for use with Charge-Transfer Devices and the Design of Functional Logic Arrays", IEEE Journal of Solid-State Circuits, Vol. SC-7.

ZIMMERMAN, T. A., 1975, "The Digital Approach to Charge Coupled Device Signal Processing", IEEE Advanced Solid State Components for Signal Processing, Circuits and Systems Conference.

ZIMMERMAN, T. A., and BARBE, D. F., 1977, "A New Role for Charge-Coupled Devices: Digital Signal Processing", Electronics.

$$S1 = A+B+G$$

$$S2 = ABG$$

$$X = AB+AG+BG$$
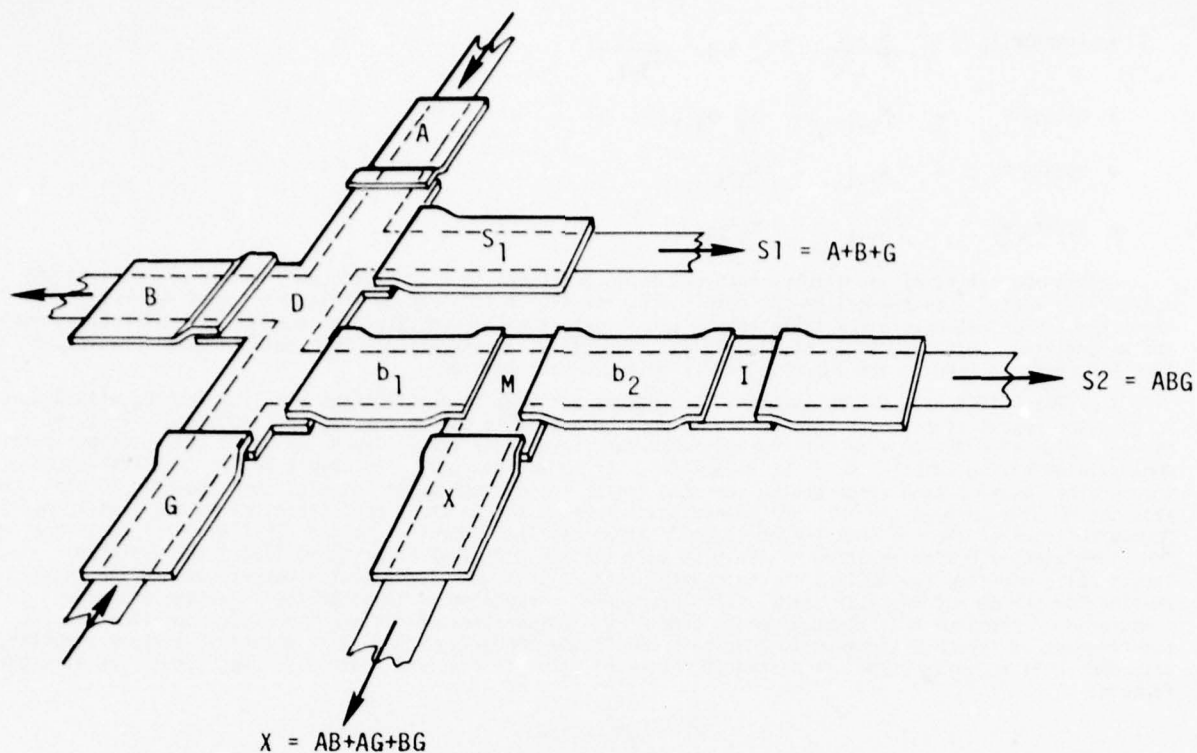
Fig.1 A three input digital charge coupled logic gate that simultaneously implements both the AND and OR functions and a combinational logic term

$$S1 = A\overline{B}+\overline{A}B$$

$$C = AB$$
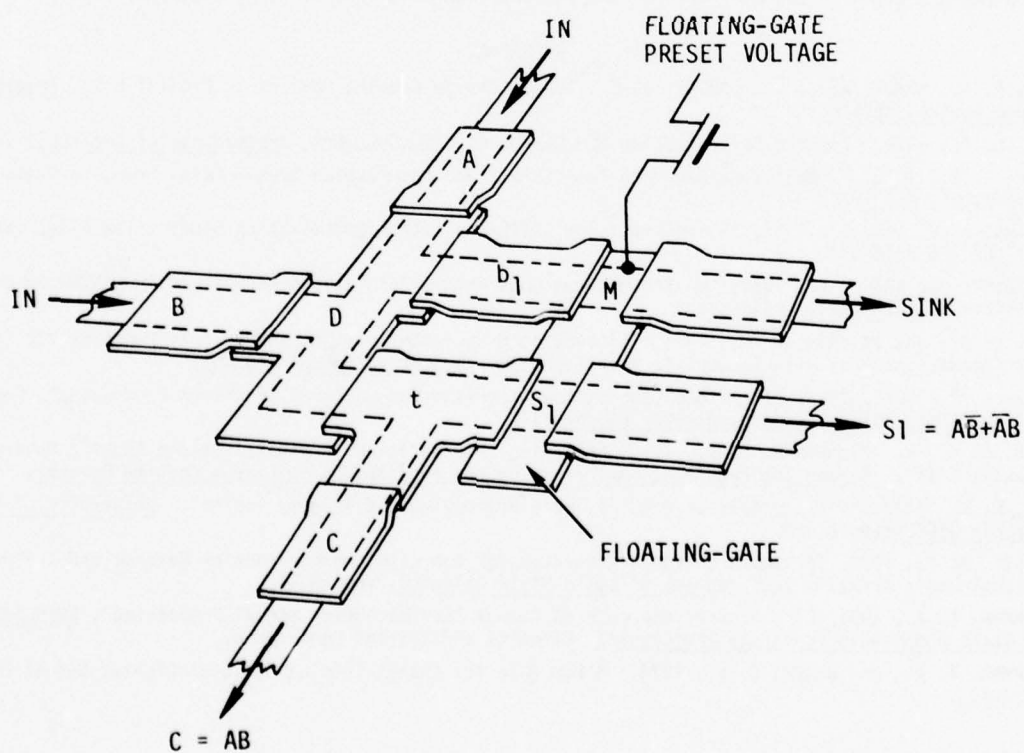
Fig.2 A digital charge coupled logic half-adder or exclusive-OR gate, showing the location of the floating gate
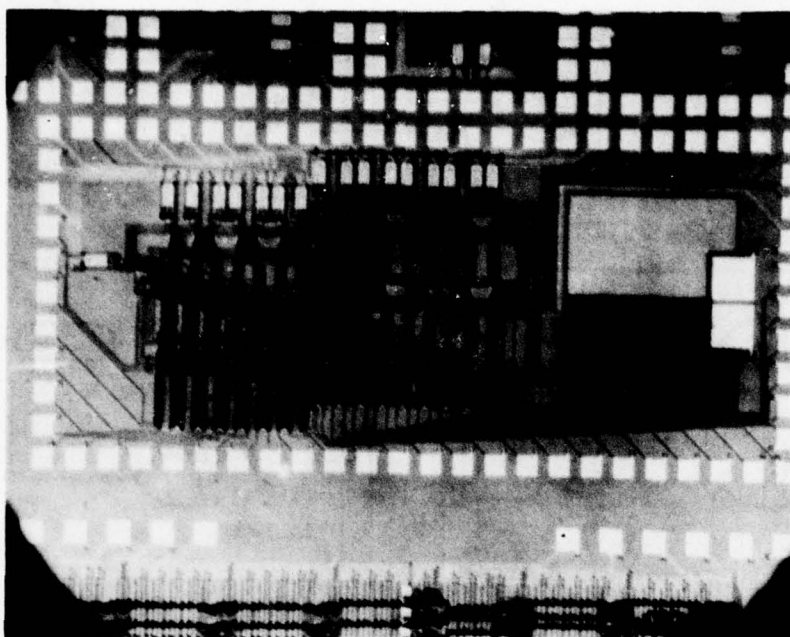
Fig.3 Photograph of processed DCCL 2-word, 16-bit, parallel binary full-adder array
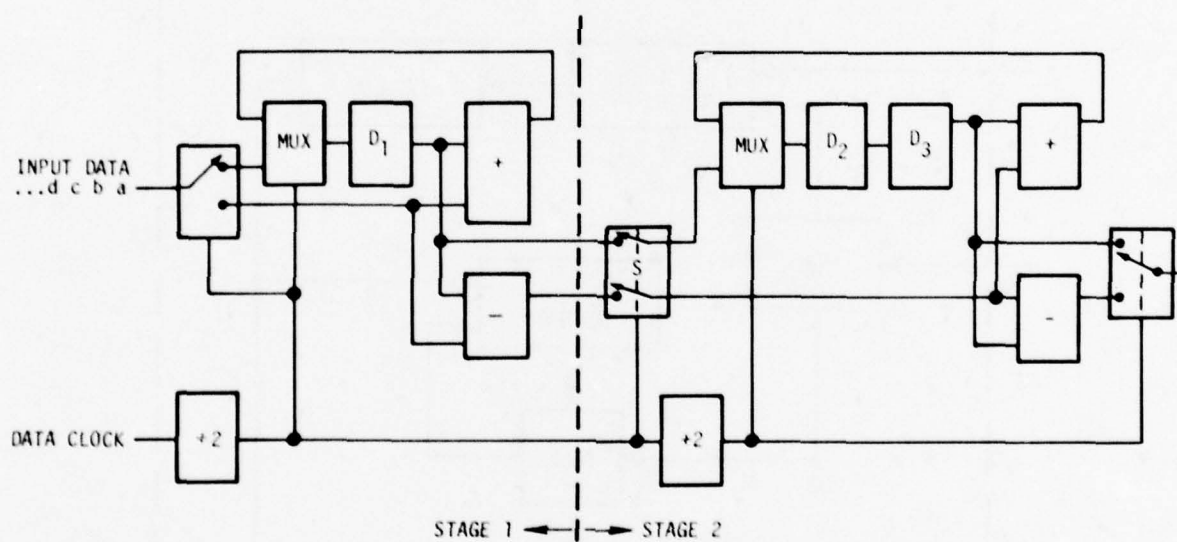
Fig.4 Basic pipeline hadamard transform stage
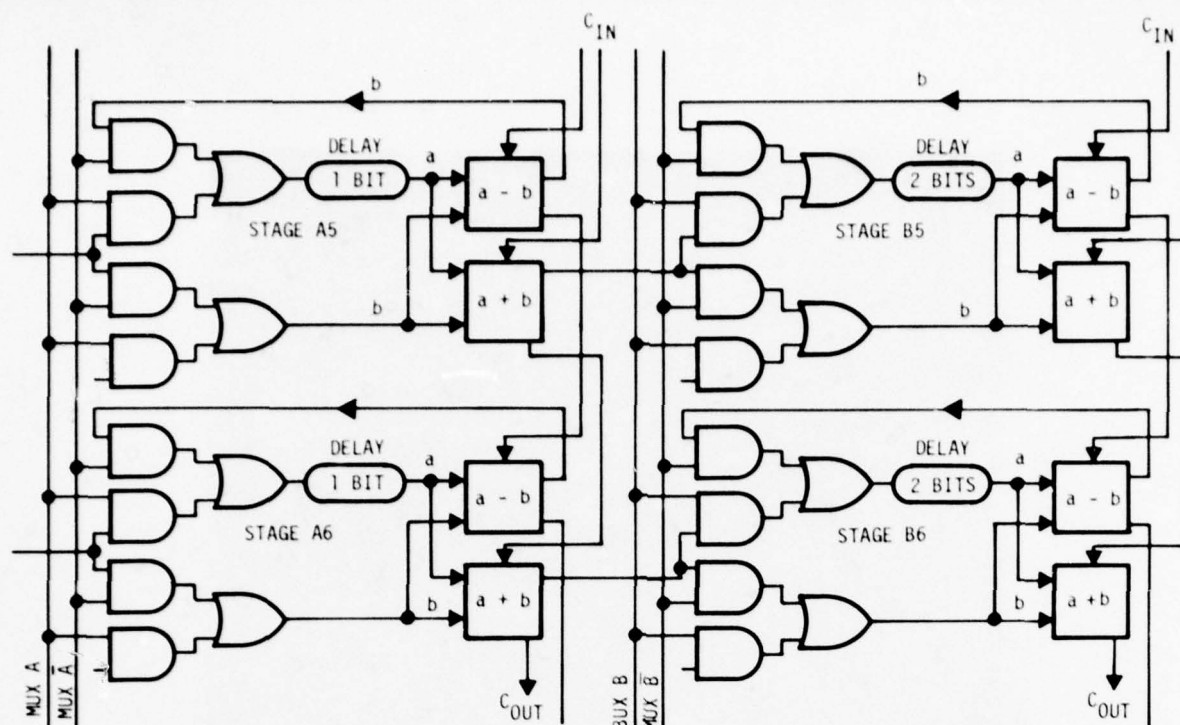
Fig.5   Four typical FHT cells
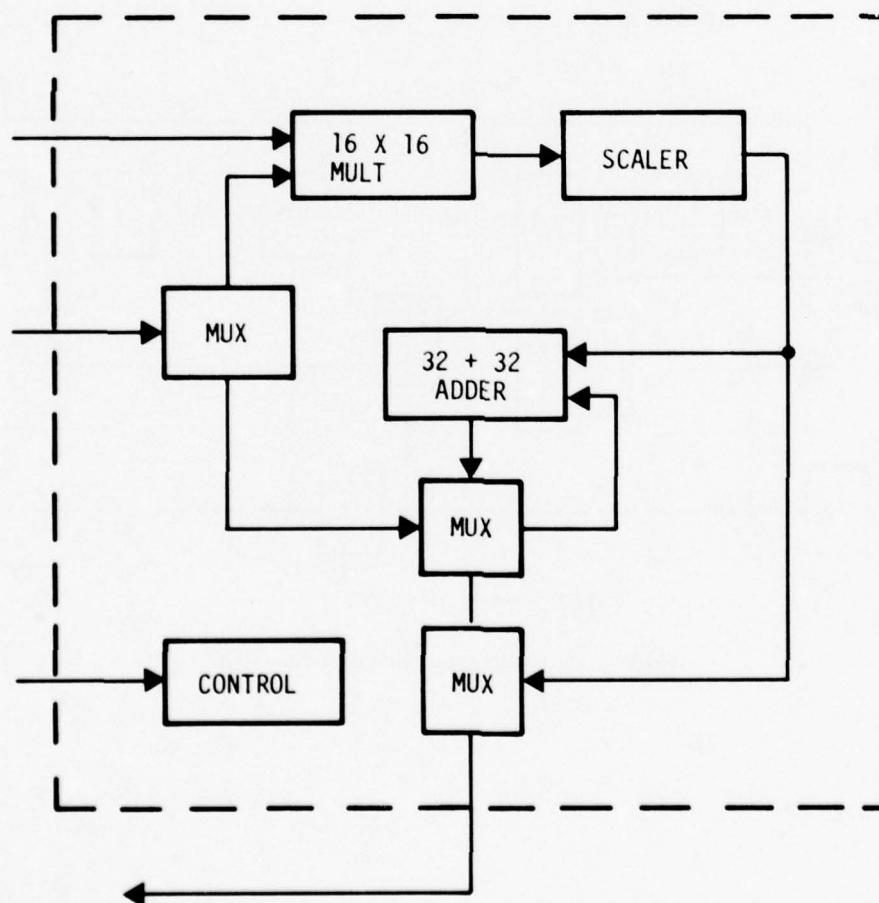
Fig.6   DCCL arithmetic unit chip
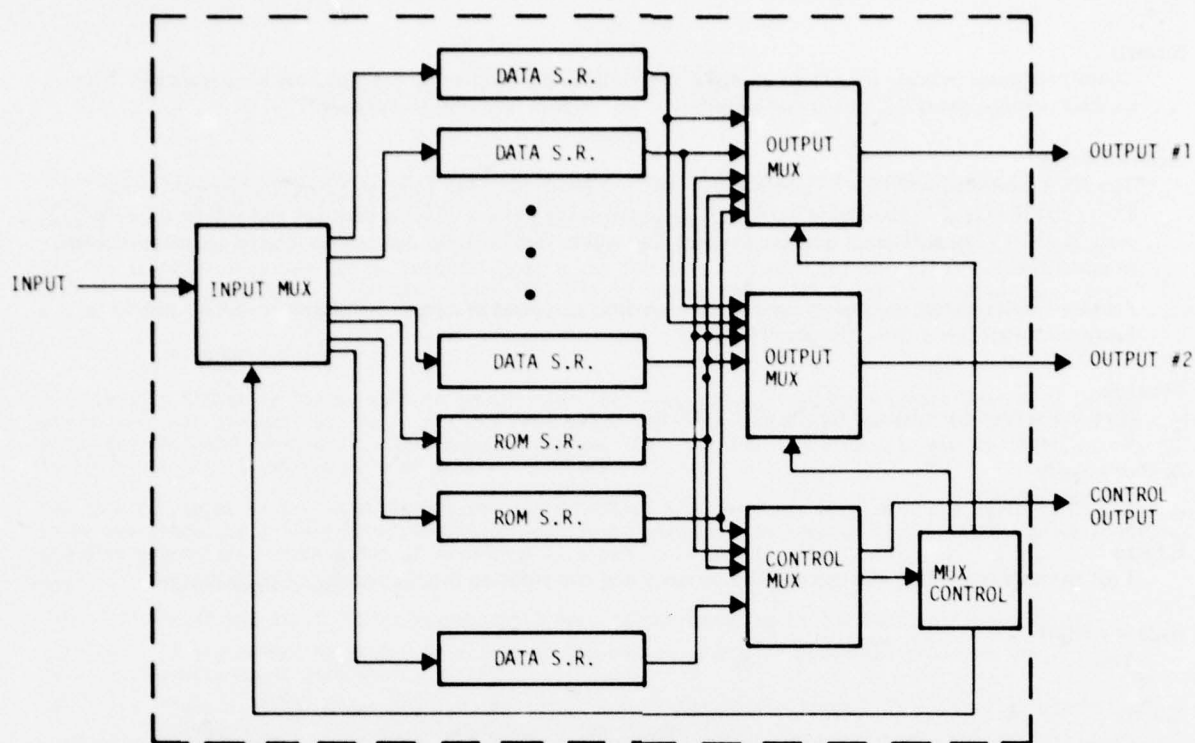
Fig.7   DCCL controller chip

$$e_m = e_{m-1} k_m + Df_{m-1}$$
$$f_m = -k_m e_{m-1} + Df_{m-1}$$
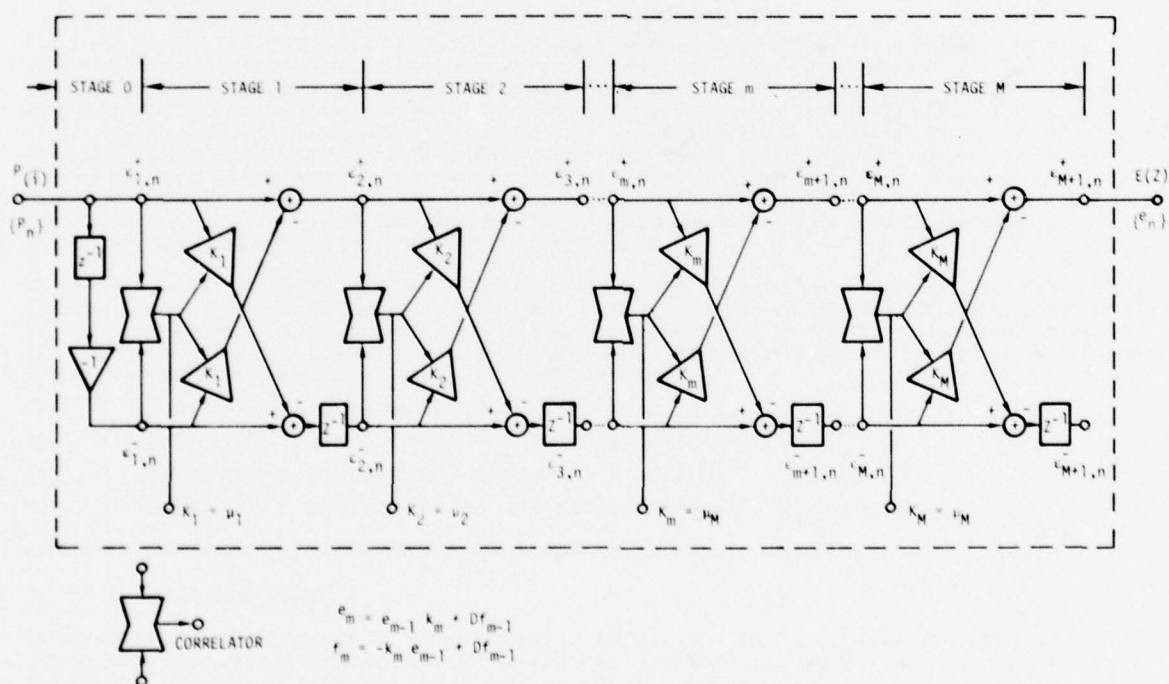
Fig.8   Itakura analyzer

DISCUSSION

Roberts

Could you please indicate the extent to which the complex processors you describe have been realized? What packing density, speed and power consumption do you expect from this technology?

Author's Reply

The 8 x 8 multiplier and 16 x 16 adder arrays are being produced. The 16 x 16, 32 + 32 adder/subtractor arrays and the (16 x 1) Hadamard Transform chips are in the process of being digitized and will be produced early in 1978. The arithmetic unit and Control chip will be designed and digitized in the spring and produced in mid 1978.

Packing density, speed and power consumption are fully discussed in a paper published by IEEE Solid State Section Journal this month (October 1977).

Tournois

How much time do you need for the FFT 1000 points you described?

Author's Reply

200 nanoseconds.

E.Stern

I got the impression that you can do both memory and computation functions using CCD. Is this so?

Author's Reply

Yes.

E.Stern

Could you contrast your technology with conventional technology TTL, especially in terms of power consumption?

Author's Reply

Power consumption is much less than TTL.

CHARGE COUPLED DEVICES WITH SIMPLIFIED DRIVE REQUIREMENTS

J.N. Gooding
Allen Clark Research Centre,
The Plessey Company Limited,
Caswell, Towcester, Northamptonshire, NN12 8EQ, England

SUMMARY

The capabilities of Charge Coupled Devices make them attractive for a wide range of signal processing applications. However, their advantages are diminished if they need to be surrounded by a significant amount of support circuitry. This paper discusses to what extent this peripheral circuitry can be simplified if not eliminated. Two delay line designs are discussed in detail. In an audio delay line all the support circuitry can be integrated onto the CCD chip. Power considerations prevent this being done at video frequencies. However, it is possible to reduce the video delay line drive requirements to a single phase clock drive - all other required waveforms, etc. being generated on chip.

1. INTRODUCTION

Our aim with CCD delay lines has been to engineer them into attractive, easy-to-use devices. In this way we can make the most of their inherent cost, size and power advantages, and allow these advantages to percolate to the system level. This paper discusses the techniques necessary to achieve this and any implications that these may have on device performance.

The main complaint of system designers who are potential CCD users has been that the devices have needed to be surrounded by a significant amount of additional circuitry to provide the necessary drive waveforms and bias potentials. Such a set of waveforms is shown in Figure 1. This shows

(a)   Three-phase clock drive

(b)   'Fill and spill' pulse necessary to perform the sampling of the input analogue signal

(c)   'Reset pulse' necessary to reset the CCD output diffusion potential prior to the emergence of each charge packet

(d)   'Sample pulse' necessary for the output sample and hold circuit. This circuit reduces the amount of clock signal present in the output signal.

In addition to these, it has been necessary to provide several biasses for input gates, output gates, substrate bias, etc. We will discuss the improvements that can be made in this situation related to CCD analogue delay lines, but these techniques can and are being applied equally to other CCD signal processing devices.

2. THE PROCESS

The first improvement that can be made is to choose a process which has built-in barriers to give directionality of charge transfer as described in (Browne, V.A., this conference). This not only reduces the number of clock waveforms which must be applied to the device but it makes the device much more tolerant to overshoots and other waveform imperfections. Having such a two-phase process then allows us to operate in the pseudo single phase mode by applying a D.C. potential to one set of electrodes and clocking the other set about this potential. On-chip MOS circuitry can then be used to integrate as much as possible of the required peripheral circuitry. For the video delay line we need a process with good high speed performance. The isoplanar, N-channel, double-level polysilicon process described in (Browne, V.A., this conference) satisfies this requirement and at the same time is an industry standard process which should help to keep manufacturing costs down.

Since CCD delay lines (both audio and video) tend to be long, involving many transfers, buried channel (rather than surface channel) technology is used. This gives superior transfer efficiency particularly at video frequencies. A further advantage is that the implant used to fabricate the buried channel is also available for use in transistors. By permuting this implant with the CCD barrier implant transistors having four different threshold voltages (V_T) are available:

(a)   polysilicon 1, no implants, $V_T \approx +1.25$ volts

(b)   polysilicon 1, with buried channel implant, $V_T \approx -8$ volts

(c)   polysilicon 2, with barrier implant, $V_T \approx +3.5$ volts

(d)   polysilicon 2, with barrier and buried channel implants, $V_T \approx -3.5$ volts

With the exception of (c) all of these prove useful to the circuit designer.

3. DELAY LINE DESIGN

It was concluded in (Browne, V.A., this conference) that for buried channel CCDs, two-phase or pseudo single-phase operation offers the same charge handling capability and signal fidelity. Thus the choice between the two modes of operation can be made on ease of drive grounds.

Two-phase operation appears at first sight to give a power consumption of less than that for the pseudo uniphase mode. A 15-volt amplitude clock is required to operate the buried channel CCD (in the uniphase mode) so logically only 7.5 volts amplitude should be required for the two-phase case. Since the power dissipated depends upon the square of the amplitude, the total dissipation for the two two-phase drivers should be only half of that for the uniphase driver. To achieve this a 7.5 volt power supply would be required and this would be in addition to the 15-volt supply. A requirement for two positive supplies would be in conflict with the easy-to-use philosophy. The 15-volt supply would still be required to provide the output reset potential and the supply for the output buffer stage. Use of 7.5 volt clocks

would also present problems at the input and a higher amplitude clock for the first electrode pair of the CCD would be required to overcome this. Generation of the two 7.5 volt amplitude clocks from the 15-volt supply would involve about the same power dissipation as the generation of a single 15-volt amplitude clock.

In pseudo single-phase operation the D.C. phase needs to be provided with a source whose impedance is low at the clock frequency. Decoupling by means of an on-chip capacitance is not practical since the silicon area required is about the same as that for the rest of the chip put together. Providing a low resistance source presents about the same or greater difficulty than providing a clock driver stage.

The conclusion is that if we allow an off-chip component (a capacitance of, say, $0.01\mu F$) to decouple the D.C. phase in the uniphase approach, the on-chip generation of two 7.5 volt amplitude clocks will give the same dissipation as the generation of a single-phase 15-volt clock. If an off-chip component is not used then the two-phase solution will be the minimum power due to the extra power involved with providing the D.C. phase.

Analogue delay lines are readily divided by their operational frequency range into audio and video categories. We have chips (adopting simplified drive principals) of both types now in processing. Delay line test chips have already been fabricated on the silicon gate process and evaluated.

3.1. Audio Delay Line

At audio frequencies, on-chip MOS circuitry is perfectly capable of performing all the required peripheral functions. Figure 2 is a block diagram of our audio delay line which is a quad 512-stage device. The chip includes circuitry for

(a)   Clock generation

(b)   Clock drive

(c)   Input bias

(d)   Output buffer and sample and hold

The signal delay can be controlled either by an external TTL clock, or by an on-chip voltage controlled oscillator. In the latter mode of operation a TTL compatible clock pulse at the internally generated clock frequency is available as an output. In applications where a number of delay line chips are to be cascaded or for other reasons need to be synchronised, the delay can be programmed on one master chip which then drives all the other chips at the same frequency.

Two-phase clock drive has been selected for this chip. At audio frequencies high efficiency clock driver stages can be made but it is very difficult to provide a low impedance source for the D.C. phase without involving significant dissipation. Hence the choice lay between

(a)   2-phase clock with no external components

(b)   Single-phase clock with off-chip capacitance

Option (a) was selected and 12-volt amplitude clocks used rather than 7.5 volt. This was at a power penalty of only 6 mW at 200 kHz (less at lower frequencies) and it avoided the input problems mentioned in Section 3.0. At these power levels silicon area dominates the argument.

3.1.1. CCD input Techniques (Audio)

The phase referred input technique (McCaughan and Harp, 1976) was adopted since this requires no clocks in addition to the main transport clock.

3.1.2. CCD Output Techniques (Audio)

At the CCD output three separate non-overlapping time slots must be allocated in each clock cycle for

(a)   Output diode reset

(b)   Transfer of signal charge from last storage site to the output diode

(c)   Sampling by output sample and hold circuit

Since the clock generation is all on chip, and the MOS circuitry has speed in hand, it is a straight-forward job to provide a clock sequence which includes pulses for (a) and (c) with the correct time relation to the falling edge of ϕ, which controls (b). These waveforms are shown in Figure 3 together with the gating circuitry necessary to generate them.

3.1.3. Audio Delay Line Performance

The performance of the device is summarised in the following target specification (quoted at $25^{\circ}C$ with $V_{DD} = 15$ volts).

PARAMETER		MIN.	TYP.	MAX.	UNITS	CONDITIONS
Per Channel:						
Audio Delay	min.	2.5	2.5	2.5	mS	f_c = 200 kHz
	max.		50		mS	20% of signal range allocated to leakage
Maximum Signal			1		Volt	Peak-to-peak
Dynamic Range		60			dB	Over bandwidth 0 to $f_{c/2}$
Total Harmonic Distortion				-50	dB	For IV peak-to-peak signal
Through Gain		-1.0		+1.0	dB	
Overall Transfer Inefficiency			3		%	
Input Impedance			100		kΩ	Bias circuit impedance
Output Impedance			5		kΩ	
Per Chip:						
Clock Frequency		5		200	kHz	(Clock Mode High: Ext. clock)
Power Consumption			100		mW	(Clock Mode Low: Int. clock)

3.2.   Video Delay Line

Here we must consider how many of the audio delay line features can be implemented at video frequencies and indeed, to what extent this is desirable. Video delay lines are required to have a pass band from D.C. (or about 100 Hz) up to about 5 MHz. This means that the CCD clock frequency must be greater than 10 MHz and more like 15 MHz in order to make input anti-aliasing filtering practical. This frequency is close to the useful limit for MOS circuitry but we are aided by having the N-channel isoplanar process and in particular its self-aligned diffusions.

The available circuitry does in fact have the speed capability to provide all the required on-chip functions at frequencies of up to about 20 MHz, but in most cases power considerations prohibit us from exploiting it to the full. This is particularly true in the case of clock drivers. As we approach the MOS useful frequency limit clock driver circuits become less efficient. This is especially true if the clock waveform is required to pull up to within, say, 0.5 volt of the V_{DD} rail.

Taking as a practical example a 1000-stage delay line using a 50µ channel width and designed to operate at 20 MHz at 125°C we find:

(a)   If 100% efficient drivers were available they would dissipate about 450 mW (uniphase 15-volt amplitude clock or two 7.5 volts clocks derived from the 15-volt rail)

(b)   A practical uniphase clock generator would have a worse case dissipation of about three watts. This mode of operation would also require an off-chip decoupling capacitance for the D.C. phase.

(c)   A practical circuit to provide two phases at about nine volts amplitude would have a worse case dissipation of about two watts.

On-chip C-MOS circuitry would obviously be useful in this respect. This would however be a non-standard process, be more complex than the isoplanar polysilicon one currently used and be rather limited in application. Thus it is not really a very attractive solution to this problem. The silicon area required for the above driver(s) would be large, maybe half the area of the CCD which it had to drive but this in itself would not be prohibitive.

Figure 4 shows the effect of on-chip dissipation on CCD storage time. This is plotted for 25°C ambient and for the situation when 10% of the storage well capacity is allocated to leakage charge. It can readily be seen that, if we are not to significantly degrade the storage time, the on-chip dissipation must be limited to the 200 mW region. These figures relate to a 16-pin ceramic package in still air. A 24-pin ceramic package could dissipate about 50% more power or a 16-pin plastic package about 40% less power for the same degradation. Any requirement for improved cooling would conflict with the easy to drive philosophy.

We can conclude that for most video delay line applications it is necessary to provide some degree of thermal isolation between the CCD and its clock driver circuit, which in practice means that they must be housed in separate packages.

We can, however, integrate all the other peripheral circuitry, leaving the user to provide only a single-phase 15-volt amplitude clock and a decoupling capacitance for the D.C. phase. All the additional clocks associated with the CCD input and output have to drive only relatively small capacitances and thus these functions can be performed by on-chip circuitry without causing undue heating of the chip.

3.2.1. CCD Input Techniques (Video)

The phase referred input technique is obviously attractive due to its simple drive requirements. We have found, however, that it is only useful up to a few MHz. It involves a spilling of charge from the metering well back to the input diode during part of the risetime of the clock waveform. If the risetime is too short insufficient spilling takes place and distortion of the transfer characteristic results. In order to avoid this problem the risetime of the clock waveform must not be less than about 70 nS at 25°C (about 140 nS at 125°C). This limits the technique to frequencies below about 3 MHz or possibly 5 MHz if an asymmetrical, non-unity mark to space ratio waveform is used.

The 'fill and spill' technique also requires a finite spill time which is about 10 nS at 25°C (20 nS at 125°C). Hence, it is useful when used in conjunction with on-chip pulse generation up to about 10 MHz, although again this could be slightly extended by tailoring the clock waveform.

The 'floating diffusion' (or charge injection) technique is suitable for operation with on-chip generated clocks up to 20 MHz. Figure 5(a) shows the structure of the CCD input and 5(b) the waveforms necessary to drive it. It is essential that the input sampling clock ϕ_S does not overlap on either edge the clock waveform applied to the electrode pair following V_{R1}. The only practical way of achieving this is to apply a delayed clock (ϕ_D) rather than the main clock to this electrode pair. Then by use of gating we can produce two waveforms ϕ_S and ϕ_D which are completely underlapping. The charge packet will reside under the ϕ_D electrode pair and the following ϕ_D electrode pair for less time than it resides under the other CCD electrodes and in this way make up the delay between ϕ and ϕ_D. The waveforms of Figure 5(b) can be generated by the gating shown in Figure 5(c), although this is not the only means of doing it.

3.2.2. CCD Output Techniques (Video)

As pointed out for the audio delay line in Section 3.1.2., three separate non-overlapping time slots must be allocated in each clock cycle for the correct operation of the CCD output and sample and hold circuit. This can be done by first generating two non-overlapping clock waveforms in an identical manner to that performed for the video delay line input circuit. The two waveforms (ϕ_D and ϕ_S) can now be used for the output diode reset and the sample-and-hold sample pulse respectively. It remains to delay the time at which the signal charge is transferred from the last storage well to the output diode so that this occurs after the reset transistor is turned off, but before the sample transistor is turned on. Two things are done to achieve this delay:

(a)   The last electrode pair is split electrically and ϕ_D is used to drive the storage electrode (polysilicon 1) whilst the normal ϕ is still used to drive the barrier electrode (polysilicon 2).

(b)   This last storage electrode is enlarged so that signal charge is not transferred to the diode until further down the falling edge of ϕ_D than would have been the case with a standard size electrode.

A depletion mode transistor using polysilicon 1 is used for the reset transistor, and a polysilicon 2 depletion mode transistor is used for the sample switch, each being chosen because of their threshold voltage.

Figures 6(a) and (b) show the output structure and timing. The shaded parts of ϕ_S and ϕ_D in Figure 6(b) represent the 'on' periods for the reset and sample transistors.

3.2.3. Video Delay Line Performance

The performance of our dual 256-stage delay line is summarised in the following target specification (quoted at 25°C):

Parameter	Min.	Typ.	Max.	Units	Conditions
Per Channel:					
Video Delay Min.	12.8	12.8	12.8	μS	f_C = 20 MHz
Max.		40		mS	20% of signal range allocated to leakage
Maximum Signal		2		Volts	Peak-to-peak
Dynamic Range	60			dB	Over bandwidth 0 to $F_{S/2}$, peak-to-peak signal: RMS noise
Total Harmonic Distortion		-50	-40	dB	For 2 volt peak-to-peak signal
Through gain	-1.0		+1.0	dB	
Overall Transfer Inefficiency		1.5		%	
Per Chip:					
Clock Frequency			20	MHz	Up to 125°C
Power Consumption		180		mW	

4.   SUBSTRATE BIAS GENERATION

The virtues of applying a reverse bias to the substrate are:

(a)   Any overshoots on the clock (or at any other circuit point) which cause an N+
        diffusion to substrate diode to become forward biassed will generate carriers
        within the substrate. This can give rise to greatly increased leakage being
        observed in the CCD. The application of a reverse bias to the substrate will
        reduce the likelihood of any overshoot occurring.

(b)   The junction capacitance between a diffusion and the substrate depends upon $\{V_B\}^{-\frac{1}{2}}$
        where V_B is the junction reverse bias. Hence an increase in V_B will reduce the
        capacitive loading due to the junction and give a correspondingly small increase
        in the circuit speed.

(c)   It will also mean that circuit voltages are fluctuating on the less steep part of
        the capacitance/voltage curve thus reducing harmonic distortion in analogue
        circuits.

The disadvantage of substrate bias is of course the complication involved in providing it. Since the
only current that this supply needs to provide is that due to leakage current, it is possible to provide
an on-chip circuit to generate it. This is a practice adopted by several other MOS circuit manufacturers
since the problem is common to all MOS (and particularly dynamic) circuits. Such a generator has been
included in the design of our audio delay line. There is, however, a hazard involved with this method of
bias generation. The bias generation circuit itself will inject carriers into the substrate and steps
have been taken in the design to isolate these from the CCD. It remains to be seen whether this is 100%
effective.

5.   CONCLUSIONS

By the means described we believe it is possible to reduce the support circuitry required for CCD delay
lines to a minimum without compromising the device performance. The availability of the next generation
of engineered CCD products similar to that shown in Figure 7, should result in a much more universal
acceptance of the devices than has hitherto been seen.

6.   ACKNOWLEDGEMENT

The author wishes to thank J.F. Dickson and W.D. Pritchard for useful discussions during the preparation
of this paper, and P.S. Paddan for carrying out relevant experimental measurements.

7.   REFERENCES

BROWNE, V.A., this conference, "Performance Limitations of Two Phase CCDs".

HARP, G. and McCAUGHAN, D., December, 1976, "Phase-Referred Input: A simple new linear CCD input
method", Electronics Letters, Vol. 12, No. 25.
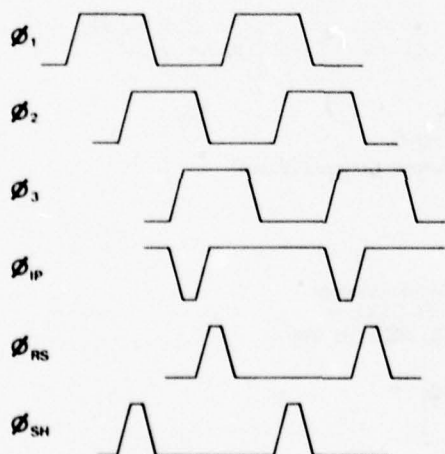
Fig 1
3 Phase Transport and I/O Clocks

Fig 2
Quad 512 Bit CCD Audio Delay Line
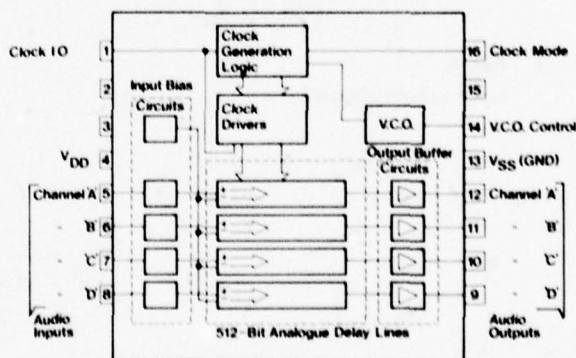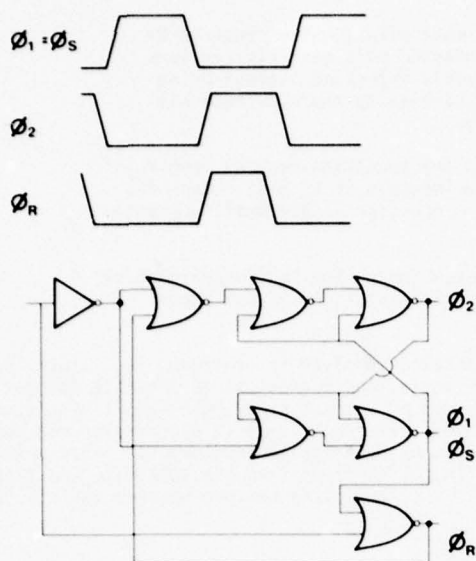
Fig 3
Output Scheme (Audio)
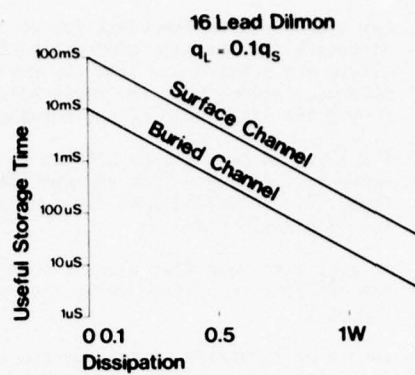
Fig 4
Effect of Dissipation on Storage Time

16 Lead Dilmon
$q_L - 0.1q_S$
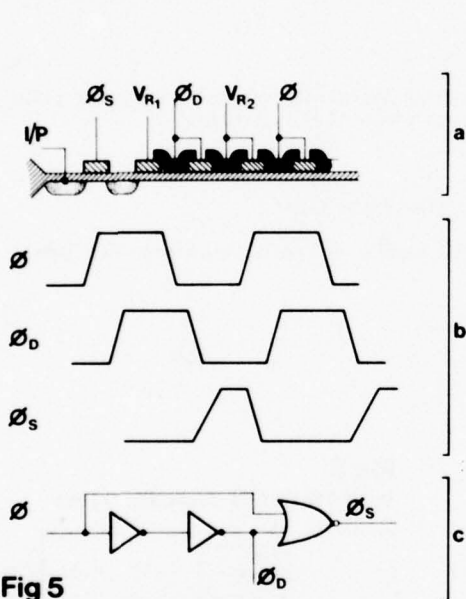
Fig 5
Input Scheme (Video)
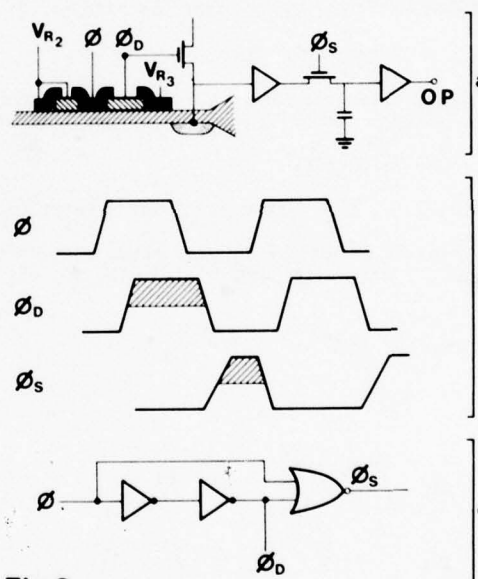
Fig 6
Output Scheme (Video)
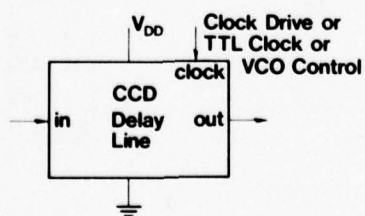
Fig 7 Simplified System

DISCUSSION

Maerfeld

Le tableau de votre papier concernant les caractéristiques des lignes vidéo sont elles des caractéristiques mesurées ou objectives?

Author's Reply

The figures are target specifications and are not proven. But we have made test lines and we are confident.

SAW FILTER APPLICATION FOR PHASED ARRAY RADAR

W. Sander
Forschungsinstitut für Funk und Mathematik
Königstraße 2, D-5307 Wachtberg-Werthhoven, W.Germany

and

W.H. Haydl
Institut für Angewandte Festkörperphysik der Fraunhofer-Gesellschaft
Eckerstraße 4, D-7800 Freiburg, W.Germany

SUMMARY

The Electronically Steerable Radar (ELRA) is an experimental phased-array system consisting of separate transmitting and receiving arrays, several coherent and incoherent signal processing units, and computers for antenna control, data handling, and the display system. The design was aimed at a versatile instrument for proving modern radar proceedings, signal process-ing and data handling techniques, incorporating multiple beam and multi-function capability [1].

This paper deals with the design, development, and fabrication of two types of SAW filters for the IF-amplifier of the re-ceiving array. Compared to conventional filters with lumped elements, these filters have some important merits. Their use represents one of the few applications where high-quality mass-produced SAW devices have been applied to improve the system performance.

1. INTRODUCTION

There are some peculiarities of the array system which influence the design of the filtering devices in the IF-amplifier.

- The transmitting array serially radiates up to six nearly rectangular pulses, each of them in a different direction. The pulse length can vary from pulse to pulse, and depends on the actual task of the radar. A length of 10 μs is used for un-detected target search resulting in a relatively large resolution cell. For target tracking the pulse length is reduced to 2 μs giving a better range resolution. The receiving array should be able to scan the six directions in which the trans-mitter has been sending and should be able to simultaneously match its bandwidth to the transmitted pulse length in order to optimize the signal-to-noise ratio.

- The receiving array consists of many identical receiver chains. At the moment we are using 200 modules, and the next stage will comprise 800 modules. With such a high number, the stability of the single receiver is an absolute demand. Alterations of parameters will gradually worsen the system performance which can only be recovered by permanent main-tenance. Though we use an automatic system for measuring and compensating relative phase drifts, there are other parameters, e.g. the frequency response, which cannot be controlled permanently and therefore may not have tempera-ture drifts or ageing effects.

- For special signal processing applications the actual frequency response is not as important as its reproducibility, both in amplitude and phase. The frequency behaviour may differ from the optimum but should be equal in all receiving channels in order to avoid the decorrelation of wideband signals. This is a very tight demand which causes many difficulties in realizing a filter.

2. DERIVATION OF THE FILTER SPECIFICATIONS

The aspects of chapter 1 have influenced the design of the receiver module (figure 1). A low-noise preamplifier, an image-rejection mixer, and a low-noise IF-preamplifier form the S-band front end. It is followed by two IF-amplifiers with different bandwidths. The first has a high bandwidth matched to the short pulses, the second has a low bandwidth matched to the longer pulses. The latter can be bypassed when receiving short pulses making it possible to change the receiver bandwidth very rapidly according to the actual pulse length. A synchronous detector delivers the in-phase and quadrature components of the signal converted to the video band allowing coherent signal processing. The phase shifter is located in the reference path of the detector and can be rapidly switched between six predetermined phase values belonging to those directions in which the transmitter has been radiating. The combination of phase shift change, bandwidth switching and appropriate sampling of the outputs represents a time-multiplex method for generating six independently steerable beams [2]. This procedure has no loss of signal-to-noise ratio as this has been previously optimized by the filters. In comparison with this, the usual arrangement of the phase shifter in the RF-stage would require a higher bandwidth matched to the switching time interval which is shorter than the pulse length. A filter matched to the signal bandwidth would deteriorate the signal by its settling time.

The most efficient filter for discriminating between white gaussian noise and the desired signal echoes is a matched filter whose frequency response is the complex conjugate of the transmitted spectrum. Thus, the matched filter for the rectangular 10 μs-pulse should have a $\sin X/X$-type frequency characteristic with a 3 dB-bandwidth of about 88 kHz, the first nulls being 100 kHz apart from the center frequency of 30 MHz. This filter with a relative bandwidth of about 0.3 % is very difficult to realize by conventional LC-filters. We tried resonant circuits with several undercritically coupled bandpass filters, using high-quality ferrite materials, and succeeded to approximate the main range of the frequency spectrum rather well. But this solution had some serious disadvantages. The tuning procedure was difficult, the reproducibility of the fre-

quency characteristic insufficient, the long-term stability poor, and the sensitivity to temperature changes too high. Therefore, we gave up the conventional filter technique and decided to choose a surface acoustic wave filter.

The second pulse length of 2 μs is used in the tracking case where the estimation of target parameters is more important than the detection capability. Thus we did not strictly apply the matched filter concept requiring the same characteristic as above except for a higher 3 dB-bandwidth of about 0.44 MHz. The first 200 receivers are equipped with conventional 1 MHz-bandpass filters consisting of capacitors and etched coils which will be replaced in the next 600 receivers by a SAW filter with a rectangular bandpass of 1 MHz bandwidth, small ripples in the passband, the sidelobes 30 dB below the passband.

Both types of filters should have a moderate insertion loss, good temperature stability and, above all, a good reproducibility (e.g. deviations of the center frequency smaller than 2 kHz at 30 MHz).

3. DESIGN AND PRODUCTION OF THE SAW FILTERS

3.1 The SAW Bandpass Filters

Todays principle materials for SAW devices are lithium niobate and quartz, which have very different acoustic and thermal properties. Lithium niobate, which is about 30 times stronger piezoelectrically but more temperature sensitive than quartz, is generally used for wideband low-loss applications. Because of the much lower piezoelectric activity of quartz, a low insertion loss is only possible with narrowband filters. Since temperature stability, which in turn means stability of frequency and phase, is of prime importance in our phased array radar application, our choice was ST-quartz, which is 42.5° rotated Y-cut, X-propagating.

The filters consist of an input and an output transducer (reciprocal, except for impedance level) on a polished surface of the ST-quartz substrate. Transducers are fabricated by evaporating aluminium and defining the electrode pattern by standard photolithography. For the filters described in this paper, we used very simple transducers: one wideband transducer with few electrodes and one narrowband transducer which almost entirely determines the frequency characteristic of the filter. Depending on their design, transducers are capable of operating at odd harmonic frequencies [3,4]. A transducer operating at the third harmonic, requires fewer and larger electrodes [5]. This is very often advantageous for the lithographic work because the yield is vastly improved. On the other hand, fewer electrodes may cause the transducer impedance level to be too high, causing matching problems and thus high insertion loss.

3.2 The Wideband Filter

In order to obtain a rectangular-type passband, it is necessary to use at least one weighted transducer. We chose a $\sin X / X$ weighting function with a total of six sidelobes multiplied with a truncated Kaiser function [6]. An in-house developed computer program was used to select the optimum number of sidelobes for the $\sin X / X$ function and the width of the Kaiser function. For example, increasing the number of sidelobes beyond 3 only lengthens the filter structure without gaining anything in performance. Diffraction effects, due to an increasingly smaller acoustic aperture of the sidelobes can lead to a degradation in performance. Decreasing the number of sidelobes seems attractive because it leads to a saving in substrate material, but unfortunately also to higher ripples in the passband. Narrowing the Kaiser function causes increasing sidelobe suppression and decreasing rolloff at the band edges.

The final design of the wideband filter is illustrated in figure 2. Because of the large bandwidth and the small piezoelectric coupling of ST-quartz, it was necessary to use the maximum possible acoustic aperture, which was 150 acoustic wavelengths, to obtain minimum insertion loss. In order to reduce the electrical resistance of the electrodes, the transducers were split into two parts and fed from the center. Double electrodes were used with the fundamental frequency at 30 MHz and a third harmonic at 90 MHz.

3.3 The Narrowband Filter

In order to increase the yield in manufacturing the filters, we chose to design this filter to operate at the third harmonic frequency. This allowed us to use fewer and larger electrodes as are required for fundamental frequency operation. The number of electrodes is still sufficient to obtain a tolerable insertion loss [4]. The design of this filter is illustrated in fig.3. The total filter response is given by the product of the two transducer frequency responses: $F(\omega) = T_1(\omega) \cdot T_2(\omega)$, where $T(\omega)$ are the $\sin X / X$ frequency responses of the narrowband transducer (many interdigital electrodes) and of the wideband transducer (few electrodes). Two filters are placed on a 50x19 mm substrate which are separated about 4 mm to keep the acoustic coupling between them low.

3.4 The Fabrication of the Filters

Because of the low operating frequency of 30 MHz, the minimum transducer electrode dimensions (finger width) are of the order of tens of μm. The total electrode area is however several cm^2. Our filters were thus made in a single 20:1 reduction step. The filter masks were cut on rubylith and photographed several times, each time varying the reduction factor by a small amount. Because there exists a spread in the center frequency of the filters if a number of filters are made using the same mask, several filters had to be made from each mask. The mask which yielded filters with center frequencies closest to 30 MHz was then selected. A chrome copy of this mask was made which was then used to fabricate all filters.

The bottom side of the 2 x 0.75 x 0.1 inch (50 x 19 x 2.5 mm) ST-quartz substrates was ground with a small diamond drill in order to scatter volume waves which are also excited to a small degree by the SAW transducers. The substrates were then

cleaned chemically and vacuum coated with 0.4 μm of aluminium, which 30 min. of argon discharge cleaning preceded. Shipley AZ 1350H positive photoresist was used to define the filter pattern. The aluminium was then chemically etched in warm (31°C) $16 H_3PO_4 : 2H_2O : 1 HNO_3 : 1 CH_3COOH$ acid by constantly removing the gas bubbles with a nitrogen stream. The etch rate was about 0.075 μm/min. Filters were then packaged, and the frequency measured with a GR 1710 network analyzer.

3.5 Adjustment of the Center Frequency

We have measured the frequency of several hundred narrowband filters and plotted the deviation from nominally 30.000 MHz in figure 4. A gaussian-type spread is obtained with an rms mean deviation of several kHz. The data shown illustrates the center frequencies after the aluminium transducers have been etched. We have used only one chrome mask for all filters all experimental conditions were kept constant throughout the fabrication of the filters. It is apparent that those filters made earlier, have a slightly higher center frequency. This effect is believed to be an ageing effect of the polished (and thus strained) surface. Experiments are under way to confirm this. It seems reasonable that such ageing effects should be even more severe in SAW devices than in bulk acoustic wave devices.

The frequency of SAW devices may be varied by a small degree by varying the mass loading effect of the electrodes. This is, for example, possible by etching the aluminium (in our case typically 0.4 μm thick) chemically thinner. The rate of frequency change for the narrowband filters (double electrodes, third harmonic) is given in Table 1. From the experimental data shown, it is also apparent why gold electrodes are unsuitable, even at these low frequencies: in order to obtain a frequency change of 1 kHz, the deposited gold thickness must change by only 0.001 μm, a value which is not practical. An increase in the center frequency is obtained by depositing a thin layer of insulating material with a higher acoustic velocity than ST-quartz. Such a material is Al_2O_3, aluminium oxide. Similarly, if the frequency must decrease, one must deposit a material with a lower acoustic velocity than ST-quartz such as ZnO, zink oxide. Insulating ZnO is obtained by sputtering in an argon with a few percent oxygen. Al_2O_3 may also be sputtered in argon or evaporated with an electron gun. At 30 MHz, the required film thicknesses for a frequency change of 10 kHz are about 0.1 μm or less, a very practical value. The filters described here were all adjusted to a frequency of 30 ± 0.001 MHz, by monitoring the frequency during the sputter or evaporation process.

TABLE 1

Mass Loading Effect on Narrowband Filter

Material	C_1
aluminium [*] electrodes	-0.1
gold [*] electrodes	-2.85
ZnO [*]	-0.5
Al_2O_3 [**]	+0.35

$\Delta f/f_o = C_1 h/\lambda$

h = layer thickness

[*] = double electrodes, third harmonic

[**] = continuous film on aluminium double electrode transducers at third harmonic, $h_{Al} = 0.0038 \lambda_3$

4. FILTER PERFORMANCE

4.1 Reproducibility

We made some measurements on the performance confirming the desired specifications. Figure 5 shows the frequency response of the filters around the center frequency demonstrating the fundamental difference between both types of filter. The maximum sidelobes of the rectangular wideband filter remain below 35 dB, 5 dB better than specified. Ripples in the passband do not exceed 0.5 dB.

The mainlobe of the narrowband filter agrees very well with the theoretical sin X/X -function, whereas the sidelobes differ in height and symmetry. But this effect seems irrelevant to the system performance, for

- the mainlobe of the frequency spectrum comprises most of the signal power, and

- the real transmitted pulse differs from the ideal rectangular pulse by its finite rise and decay time causing a spectrum with a reduced sidelobe level.

In order to get a critical measure of the repeatibility of the frequency response we applied the following procedure. We arranged one of the narrowband filters in the reference path and another one in the test path of an HP 8407 A network analyzer which thus directly delivered the attenuation ratio and the phase difference of both filters at its outputs. In figure 6 we compared 15 different filters with the same reference filter whose frequency characteristic in amplitude and phase has been drawn in the lower part of the diagram. The frequency range of the measurement covers only a section of the principal lobe, because the reference level must be kept sufficiently high. One can infer the cause of the inaccuracies from the form of the difference functions drawn with an enlarged scale in the upper part of the diagram. Deviations of the center frequency produce odd functions, and deviations of the bandwidth produce even functions. Other functional courses point to a mixture of both causes. These deviations must have an rms-value smaller than 1 kHz, in order to involve the measured differences. This excellent value cannot be obtained with conventional LC-filters.

Differences of attenuation and phase at the center frequency (tuned out with this measurement) are naturally larger, but of no great importance for our application as we have provided hardware and software means to compensate the differences for one frequency.

4.2 Insertion Loss, Impedance Matching and Far-Off-Selectivity

The insertion loss is relatively high but can yet be reduced by proper matching of the input and output impedances which are complex, mainly capacitive. In general, there are several combinations of matching circuits which compensate the reactive component and simultaneously transform the resistive component of the filter impedance. The matching network must have a bandwidth considerably larger than that of the filter in order to have little influence on the resulting frequency response. Otherwise it could worsen the excellent performance of the SAW filter. So it is often better to admit a small mismatch, thus increasing the bandwidth at the cost of a less reduced insertion loss which can be easily compensated in the IF-amplifier at these frequencies.

The effect of the impedance matching on the insertion loss and far-off selectivity can be demonstrated in figure 7. The diagram on the right shows the attenuation function of the rectangular filter in a wider frequency range. The passband at 30 MHz is repeated at the third harmonic of 90 MHz. The peaks near 48 and 55 MHz are caused by undesired bulk waves. The gradual increase of the base-line with higher frequencies seems to originate from electrical coupling by an unfavourable measuring arrangement. At the left, the filter is matched to the 50Ω network analyzer. The insertion loss decreases around the center frequency and increases outside this region. The unwanted, spurious responses have been considerably suppressed.

The narrowband filter shows a similar behaviour (figure 8). There are repeated passbands at 10, 30, and 90 MHz, and peaks influenced by bulk waves at about 17, 19, 48, and 54 MHz. The matching network is a high-pass type, reducing the lower frequencies more than higher ones. In figure 9, two types of low-pass matching networks have been used with a higher suppression at higher frequencies whereas the reduction of the lower spurious responses is not so good. By a combination of both types, for example, a low-pass matching network in the input of the filter and a high-pass type in the output, an equalized reduction of unwanted responses can be achieved.

Further means for increasing the far-off selectivity can be very simple. Usual resonant circuits with a low Q and therefore good stability are sufficient.

For getting a high selectivity and amplification, conventional IF-amplifiers use several resonant circuits or band filters with amplifying devices between them. With SAW filters the design of IF-amplifiers can be simplified. Since the selectivity is concentrated in the filter, the amplification can be condensed in a few high-gain, wideband devices, e.g. operational amplifiers or thick-film modules. Figure 10 shows the far-off selectivity of the 1 MHz-IF-amplifier (without the preamplifier) where all spurious responses are suppressed more than 60 dB.

5. CONCLUSIONS

The experimental results obtained with a high number of filters prove that the high quality of SAW filters developed in the laboratory can be projected to the fabrication process. SAW filters are promising alternatives for phased array radar applications. They improve not only single receiver performance by realizing almost ideally the matched filter concept but also the array performance by the excellent reproducibility of the frequency response and by the stability of their parameters.

ACKNOWLEDGEMENT

REFERENCES

[1]   Hüschelrath, G.   The ELRA Phased-Array Radar with Automatic Phase Adjustment in Practice.
      Sander, W.   AGARD Conference on New Devices, Techniques and Systems in Radar, 14 to 17 June
                               1976, The Hague, Netherlands.

[2]   Sander, W.   Mehrkeulenbildung mit elektronisch gesteuerten Antennen.
                               Symposium über Radartechnik, 13 to 15 November 1974, Munich, W.Germany

[3] Engan, H.

Excitation of Elastic Surface Waves by Spatial Harmonics of Interdigital Transducers.
IEEE Trans. on Electronic Devices, Vol. ED-16, December 1969, pp. 1014-1017.

[4] Smith, W.R.
Pendler, W.F.

Fundamental and Harmonic-Frequency Circuit-Model Analysis of Interdigital
Transducers with Arbitrary Metallization Ratios and Polarity Sequences.
IEEE Trans. on Microwave Theory and Techniques, Vol. MTT-23, November 1975,
pp. 853-864.

[5] Bristol, T.W.
Jones, W.R.
Snow, P.B.
Smith, W.R.

Applications of Double Electrodes in Acoustic Surface Wave Device Design.
IEEE Ultrasonics Symposium, 1972.

[6] Tancrell, R.H.

Analytic Design of Surface Wave Bandpass Filters.
IEEE Trans. on Sonics-Ultrasonics., Vol. SU-21, January 1974, pp. 12-22.
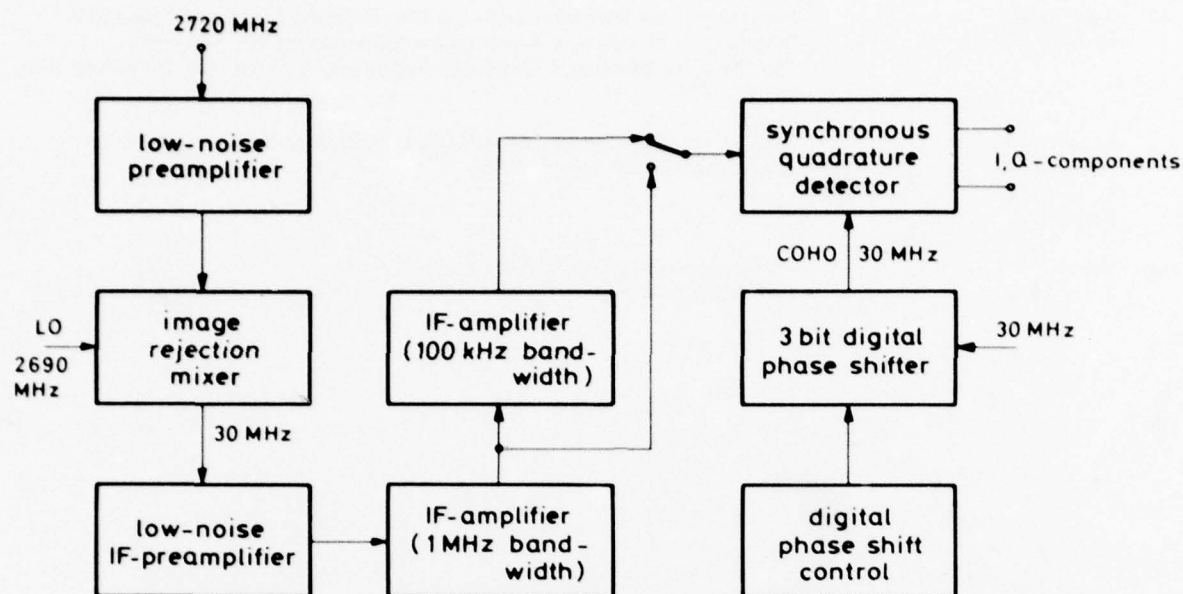
Fig. 1   Block diagram of receiver module.
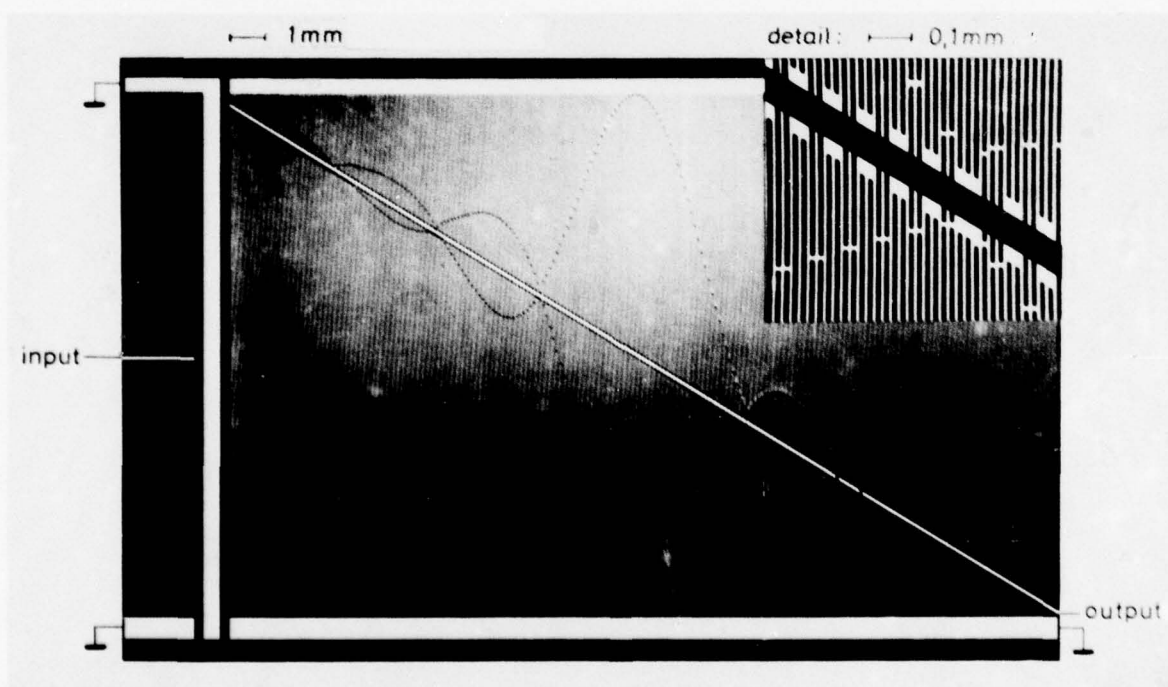
Fig. 2   1 MHz wideband SAW filter. Sin X/X -type weighting gives flat passband with -35 dB sidelobes.

input A

input B

output A

filter A

filter B

output B

— 1 mm

detail : — 0.1 mm
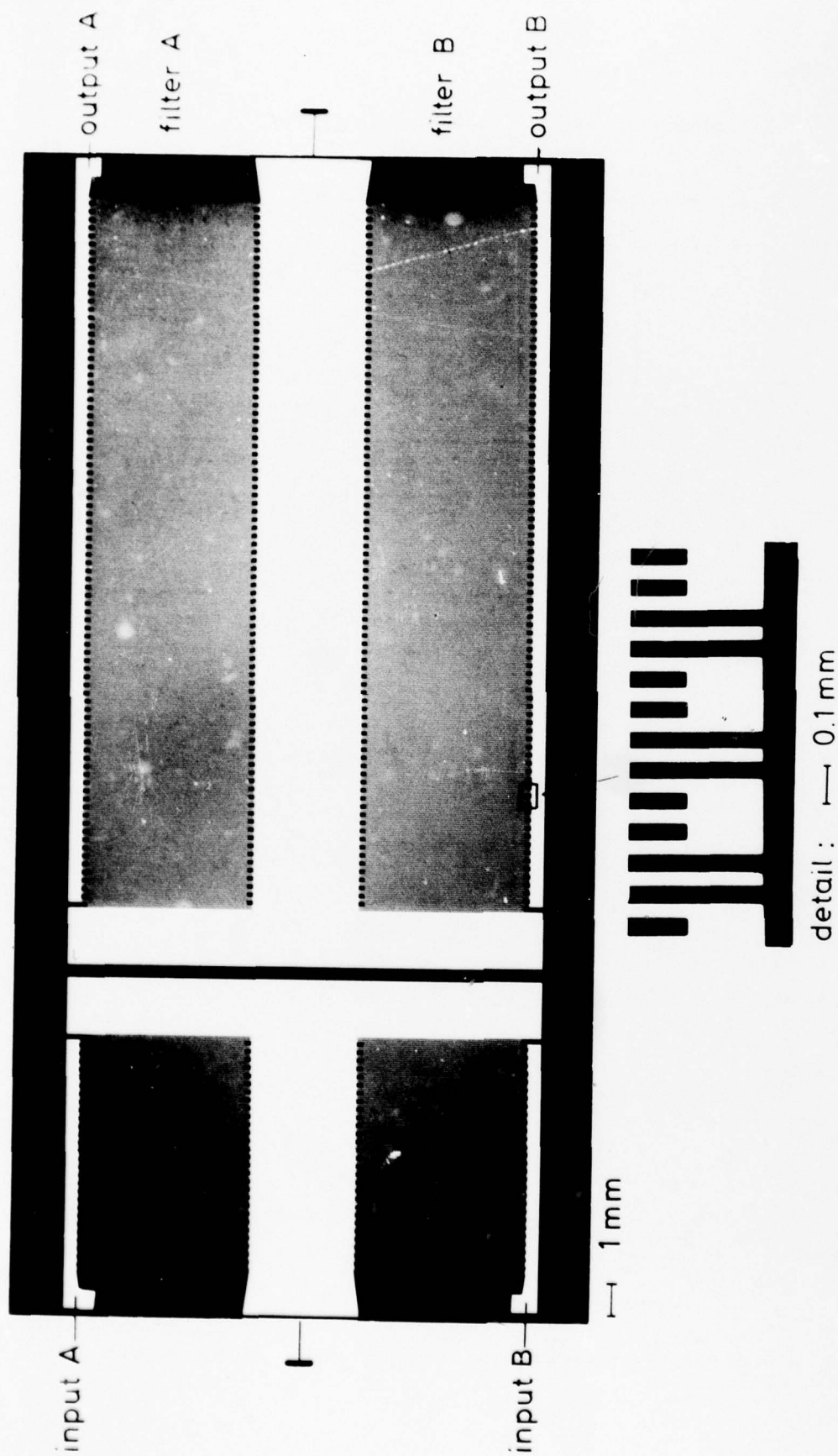
Fig. 3   0.1 MHz bandwidth sin X//X -type SAW filters. Two filters are placed on one 50 x 19 mm ST-quartz substrate. Filters are separated about 4 mm to keep the acoustic mutual coupling low.
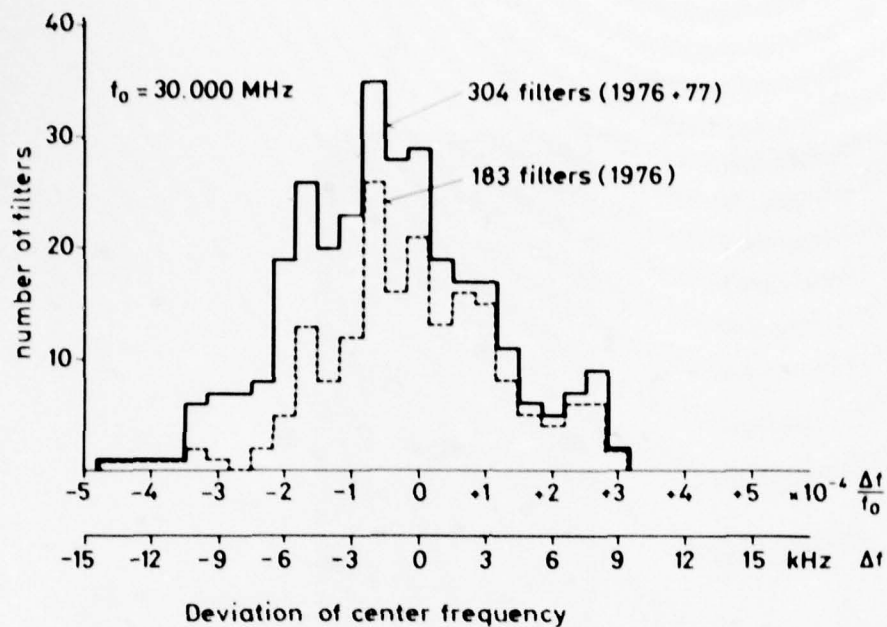
Fig. 4   Distribution of center frequency measured with 487 narrowband filters.
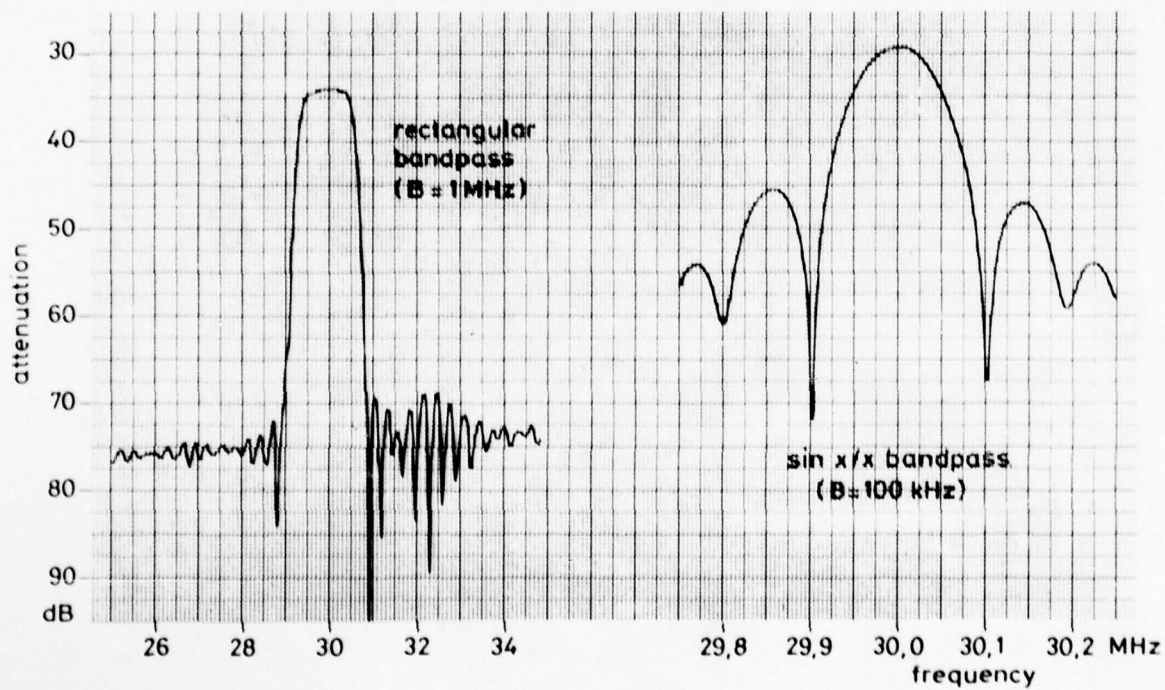
Fig. 5   Insertion loss around center frequency showing the basically different frequency responses of both SAW filters.
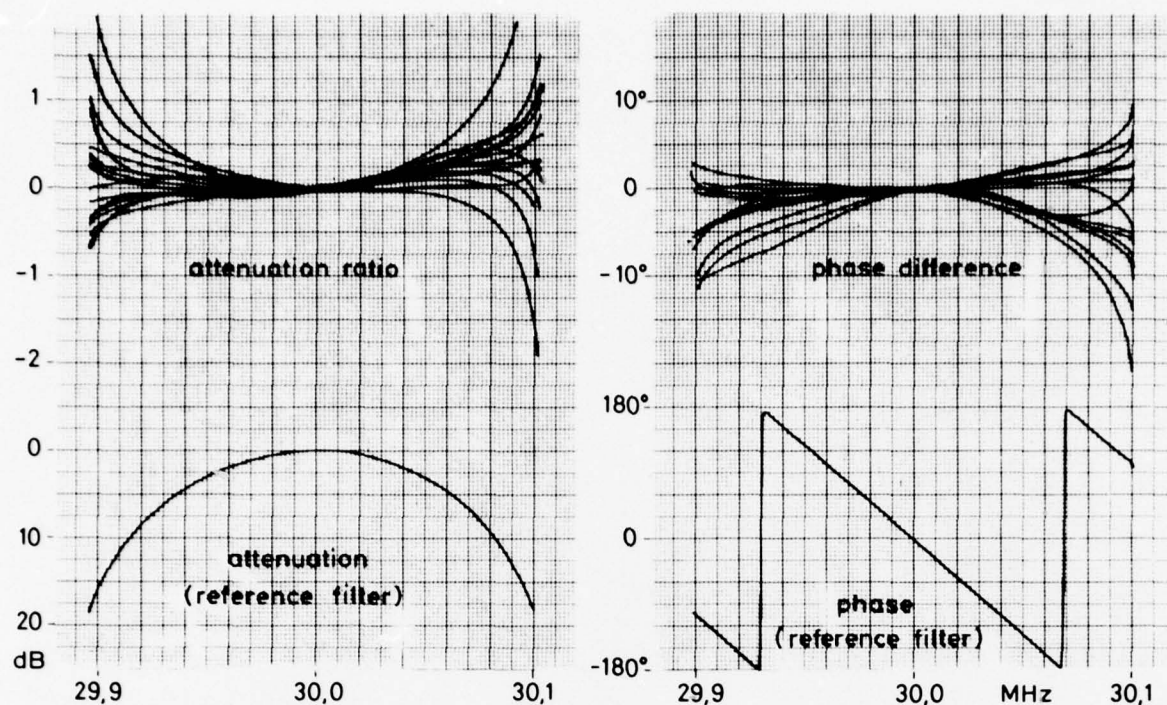
Fig. 6   Attenuation ratio and phase difference of 15 filters compared to one reference filter.
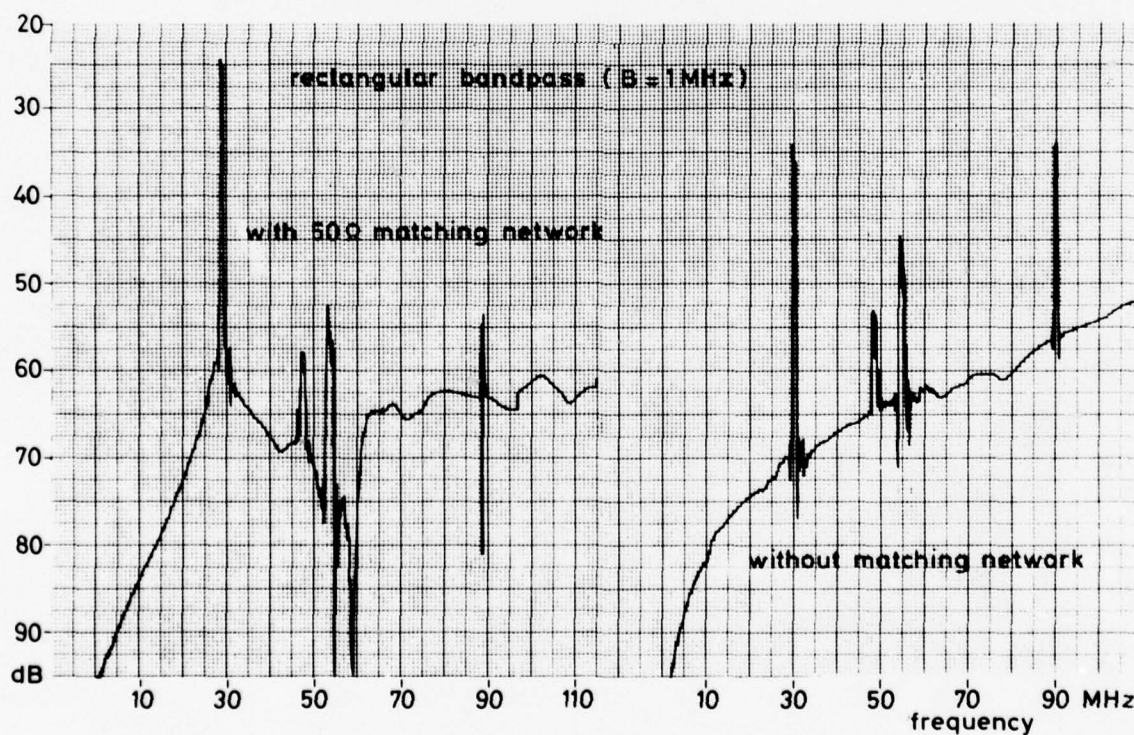
Fig. 7   Insertion loss of wideband SAW filter (right) with peaks caused by harmonic waves and bulk waves. The passband
increases, and undesired peaks decrease by matching the filter ports to the impedance of the network analyzer (left).
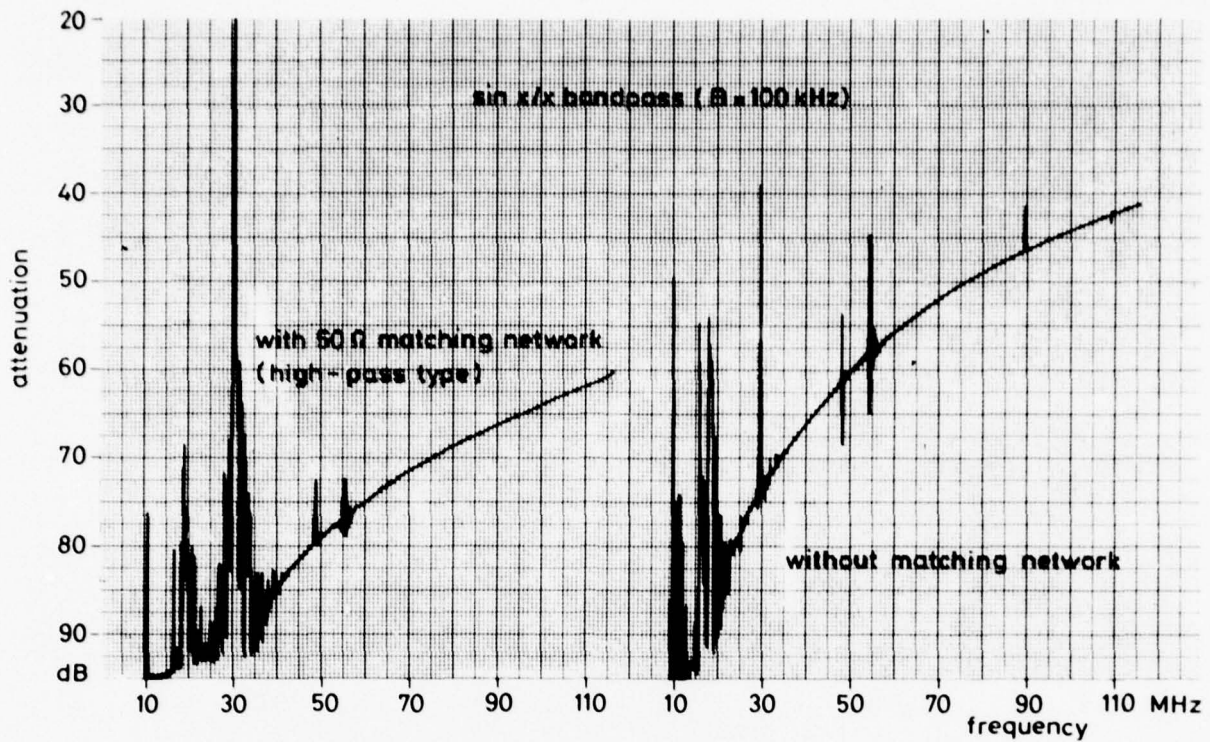
Fig. 8   Insertion loss of narrowband SAW filter before (right) and after matching (left) with high-pass networks.
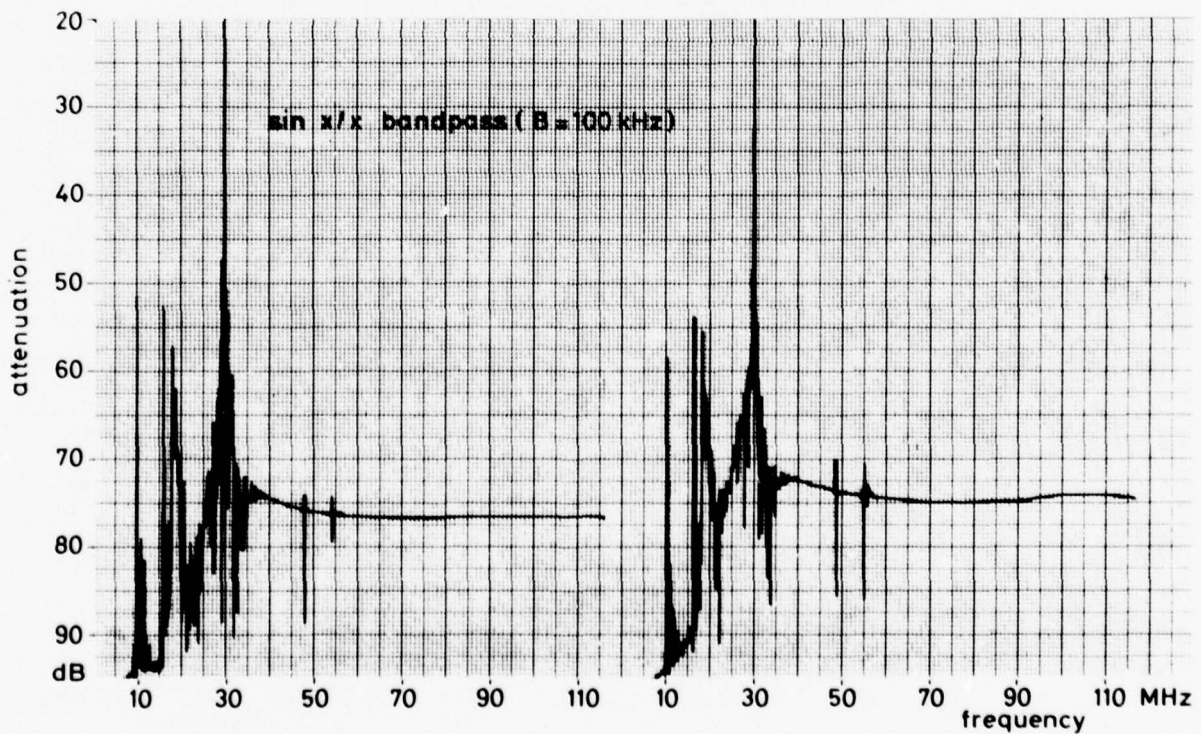
Fig. 9   Insertion loss of narrowband SAW filter matched with different types of low-pass networks.
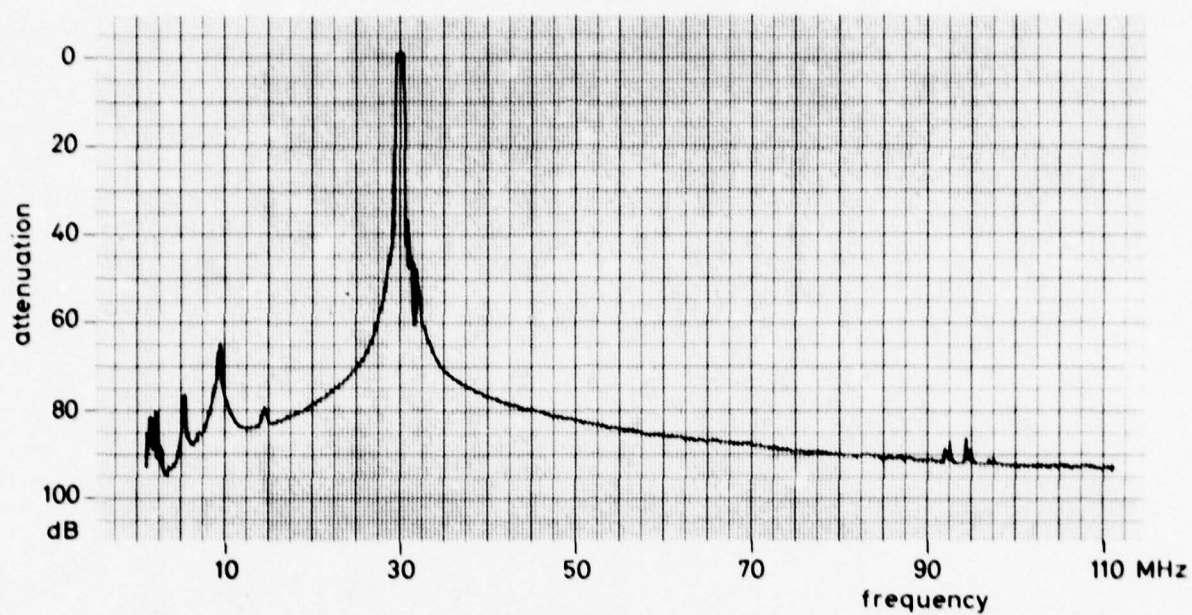
Fig. 10   Frequency response of IF-amplifier switched to the 1 MHz bandwidth with undesired responses suppressed more than 60 dB.

MODELISATION DE STRUCTURES METAL-ISOLANT-SEMICONDUCTEUR SUR Cd Hg Te

APPLICATION AUX DISPOSITIFS A TRANSFERT DE CHARGE

POUR IMAGERIE INFRAROUGE

-

M. SIRIEIX
Société Anonyme de Télécommunications
41 Rue Cantagrel
75624 PARIS CEDEX

M. GARCIA
Laboratoire d'Electronique et de Technologie
de l'Informatique - CENG -
85 Avenue des Martyrs
38041 GRENOBLE CEDEX

J. FARRE
J. SIMONNE
Laboratoire d'Automatique et d'Analyse des Systèmes
7 Avenue du Colonel ROCHE
31400 TOULOUSE

-

RESUME

L'application de dispositifs à transfert de charge à la lecture et au traitement du signal d'une mosaïque de détecteurs infrarouge permet d'envisager une nouvelle génération de senseurs dans laquelle le nombre d'éléments sensibles serait considérablement accru.

Parmi les différentes approches possibles nous avons choisi l'étude des dispositifs monolithiques réalisés sur un semiconducteur à faible "gap" : le tellurure de mercure cadmium dont l'intérêt pour la détection infrarouge dans les bandes 3-5 et 8-15 µm est bien connu.

La réalisation de cet objectif est conditionnée par certains choix dépendant de l'optimisation entre les différentes contraintes fixées par l'utilisation du dispositif. Ce sont:

- la fenêtre spectrale
- le type du matériau semiconducteur
- la température de fonctionnement
- le mode d'adressage et de lecture
- le traitement du signal

La première partie de cet exposé sera consacrée à la modélisation de la structure élémentaire de la matrice adaptée à la lecture par transfert de charge. Cet élément M.I.S détecte le rayonnement infrarouge et intègre les charges créées.

Ces deux fonctions conduisent à des limitations qui modifient la conception des systèmes infrarouge. Nous indiquerons à partir des résultats de cette modélisation les performances prévisibles pour ces dispositifs.

Dans une seconde partie, le mode d'adressage et de lecture d'une matrice bidimensionnelle sera abordé. La comparaison entre Charges Injection Devices et Charges Coupled Devices constitue l'un des choix importants dicté d'une part par la technologie et d'autre part par les propriétés du matériau.

Une discussion sur les choix possibles pour réaliser une telle matrice sera abordée compte tenu des résultats précédents.

Nous décrirons ensuite la solution retenue pour la réalisation d'une matrice 8 x 8 éléments.

I - INTRODUCTION

L'évolution des performances demandées aux systèmes d'imagerie infrarouge conduit à augmenter le nombre d'éléments sensibles placés dans le plan focal. Les technologies actuelles de barrettes de senseurs infrarouge nécessitent une connexion et un préamplificateur par élément sensible. Cette contrainte ne permet pas d'envisager des structures matricielles dépassant une soixantaine d'éléments. L'application des dispositifs à transfert de charges pour la lecture et le traitement du signal issu d'un senseur infrarouge permet la réalisation de structures bidimensionnelles à grand nombre d'éléments.

Trois types de structures sont possibles:

- l'utilisation de silicium dopé permettant une détection extrinsèque à basse température T < 77°K dans l'infrarouge, cette technique profitant de la technologie bien connue dans le silicium,

- l'utilisation de structure hybride où la détection est réalisée dans un matériau à faible largeur de bande interdite et le traitement dans le silicium. On profite alors de l'expérience technologique acquise sur les deux matériaux, la difficulté résidant dans l'hybridation,

- l'utilisation du matériau à faible largeur de bande interdite pour réaliser toutes les fonctions. La difficulté réside dans la nouveauté technique et dans la complexité des paramètres technologiques à contrôler.

Parmi les différents matériaux utilisables dans l'infrarouge proche et lointain nous avons choisi le tellurure de cadmium mercure dont l'intérêt est bien connu pour la détection infrarouge et pour l'expérience acquise depuis de nombreuses années par la SAT sur ce matériau.

La réalisation de cet objectif est conditionnée par certains choix qui ont nécessité une modélisation physique de la structure élémentaire, et, des processus de transfert compte tenu des contraintes propres à l'infrarouge.

II - MODELISATION DES STRUCTURES MIS SUR Cd Hg Te

II.1 - Principe de détection par MIS

Un senseur *MIS* est constitué par un *semiconducteur recouvert* d'une mince couche isolante sur laquelle est déposée une électrode transparente. L'application d'un potentiel sur l'électrode permet de créer une zone de dépeuplement profonde qui collectera les charges générées par l'absorption de photons dans le semiconducteur (1).

Ce processus peut se poursuivre tant que le régime permanent correspondant à la tension appliquée n'est pas atteint.

A l'issu d'un temps d'intégration, ces charges seront lues soit par injection dans le substrat soit par transfert sous une électrode voisine.

II.2 - Etude du régime transitoire

Le régime transitoire est gouverné par deux courants principaux, l'un lié à l'absorption des photons incidents, l'autre à la génération thermique des porteurs dans la zone dépeuplée (Figure 1). Cette génération thermique est caractérisée par la durée de vie en surface des porteurs minoritaires, par la profondeur de la zone de dépeuplement et par la concentration intrinsèque du matériau. On peut lui associer un temps de relaxation dans l'obscurité T_R qui doit être rendu maximal pour minimiser l'importance de la génération thermique devant la génération photonique.

II.3 - Limitation de la quantité de charges intégrable

La quantité maximum de charges que l'on peut intégrer dans la structure MIS est directement liée à la profondeur de la zone dépeuplée. En général cette profondeur est limitée par le champ de claquage de l'isolant. Dans le cas de matériaux à faible largeur de bande interdite utilisés pour la détection infrarouge, un mécanisme de génération supplémentaire produit par un effet tunnel bande à bande (2) (3) limite la quantité maximale de charges d'origine photonique que l'on peut intégrer.

II.4 - Résultats de la modélisation, Paramètres T_R, N_{max}.

Le tellurure de cadmium mercure est un alliage dont la composition détermine à une température donnée une longueur d'onde de coupure. On peut donc à l'aide d'un modèle décrivant les caractéristiques de ce matériau, étudier le comportement d'une structure MIS pour différentes longueurs d'onde de coupure et différentes températures. Un paramètre supplémentaire, la vitesse de recombinaison en surface S, a été introduit pour

traduire la qualité de l'interface isolant semiconducteur.

Sur la figure 2, nous présentons l'évolution du temps de relaxation thermique T_R en fonction de la longueur d'onde de coupure λ_c dans le cas où l'isolant est du monoxyde de silicium. Par exemple pour une longueur d'onde de coupure de 10 μm le temps de relaxation T_R ne dépasse pas 0,5 ms et peut descendre au-dessous de 0,1 ms et même jusqu'à 10 μs en pratique.

L'augmentation du temps de relaxation et de la quantité de charges maximales stockables peut être obtenue par deux moyens (voir figure 3):

- l'utilisation de diélectriques a permittivité élevée tel qu'un oxyde anodique ou l'oxyde de titane

- la diminution du dopage du matériau qui freine l'influence de l'effet tunnel. Ce résultat est directement déduit de la valeur maximale du potentiel de surface au-delà duquel le courant tunnel devient supérieur au courant de génération thermique dans la zone dépeuplée. La figure 4 montre cet effet pour deux longueurs d'ondes 5 et 10 μm.

II.5 - Modélisation du transfert de charges

Le processus de transfert de charges s'effectue par l'intermédiaire de trois mécanismes principaux (4):

- la diffusion thermique
- le champ auto induit
- le champ d'effet de bords

Différentes approches de ce problème, généralement pour évaluer l'efficacité de transfert, ont été présentées par de nombreux auteurs. La méthode que nous avons utilisée consiste en l'emploi d'un modèle simplifié permettant, compte tenu du temps d'intégration limité, de définir une géométrie de structure élémentaire compatible avec la réalisation de matrices à grand nombre d'éléments.

Le temps de transfert peut être caractérisé par une constante de temps de la forme $L^2 / 2D_p$ où L est la longueur de l'électrode et D_p le coefficient de diffusion des porteurs transférés.

En pratique, pour transférer 90% de la charge en 1 μs, la longueur de l'électrode doit être de l'ordre de 20 μm.

III - LIMITES TECHNOLOGIQUES

Les limitations technologiques proviennent des problèmes liés:

- Au matériau lui-même, qui est un composé ternaire

- Aux isolants, qui doivent être transparents dans l'infrarouge et présenter à l'interface avec le semiconducteur une faible densité de pièges.

III.1 - Le Matériau

Il doit être de bonne qualité cristalline, homogène en composition et faiblement dopé.

A l'heure actuelle des plaquettes monocristallines de 18 mm de diamètre présentent une homogénéité en longueur d'onde de coupure de ± 0,5 um pour λ_c = 10 um (6). Le dopage de type N est obtenu par compensation de l'écart à la stoechiométrie d'un matériau de type P. Les concentrations atteintes sont voisines de 10^{15} cm^{-3}.

Notons que pour les faibles largeurs de bande interdite la limite inférieure des dopages réalisables est de l'ordre de 10^{14} at cm^{-3} puisque la concentration intrinsèque est elle-même supérieure à 10^{12} cm^{-3} (λ_c = 14 μm).

III.2 - Les isolants

Il existe deux catégories d'isolants, les oxydes natifs qui se développent au détriment du substrat, et les diélectriques rapportés.

Dans le premier groupe on peut noter l'oxyde thermique difficilement controlable du fait de la décomposition du substrat à faible température (100°C) et l'oxyde anodique qui présente une bonne adhérence, une bonne reproductibilité, une transmission satisfaisante même au-delà de 10 μm et un champ de claquage de l'ordre de 10^6 V cm^{-1}.

Les isolants rapportés sont en grand nombre. Ce sont principalement le SiO, le

SiO$_2$ pour les faibles longueurs d'onde de coupure (3-5 μm), l'Al$_2$O3 et le TiO$_2$ pour toute la gamme des longueurs d'onde. Ils sont déposés en phase vapeur, ou à l'aide d'un canon à électrons, le substrat devant être maintenu à une température inférieure à 80°C.

La difficulté du choix de l'isolant réside dans le fait qu'aucun d'entre eux n'est aujourd'hui, parfaitement satisfaisant. En effet les oxydes natifs conduisent à des interfaces graduelles dont les propriétés électriques, et notamment la vitesse de recombinaison en surface, doivent être améliorées. Les oxydes rapportés induisent un phénomène d'hystérésis sur les caractéristiques capacité-tension de la structure MIS.

IV - DETECTION INFRAROUGE A L'AIDE DE SENSEURS MIS

IV.1 - Influence du flux de photons ambiants sur le temps de relaxation de la capacité MIS

Les photons absorbés par le semiconducteur créent un courant de génération qui vient s'ajouter aux courants d'obscurité précédemment décrits pour participer à la charge de la couche d'inversion.

En présence d'un flux de photons Φ_B $(\lambda_1, \lambda_2, \Omega_v)$ dépendant de la largeur spectrale (λ_1, λ_2) du filtre froid placé devant, le senseur, et, de l'angle solide de vue Ω_v, le temps de relaxation T_B est donné approximativement par la relation:

$$\frac{1}{T_B} \simeq \frac{1}{T_R} + \frac{\eta \, \Phi_B(\lambda_1, \lambda_2, \Omega_v)}{N_{max}}$$

η représente le rendement quantique.

Dans un système qui utilisera ce type de senseur, on devra choisir le temps d'intégration Ti inférieur à T_B de telle sorte que le senseur ne soit pas saturé par le fond ambiant.

Si la quantité de charges maximale Nmax que peut recevoir le senseur est sensiblement constante en fonction de la largeur de bande interdite du matériau semiconducteur, il n'en est pas de même du flux de photons ambiant qui croît fortement avec la longueur d'onde.

On peut d'ores et déjà prévoir une décroissance importante du temps d'intégration des senseurs MIS lorsque la longueur d'onde du rayonnement à détecter croît.

IV.2 - Définition d'un facteur de mérite

Pour rendre compte de la sensibilité du senseur MIS en présence des contraintes apportées par le flux de photons ambiant dans une utilisation en imagerie thermique nous définirons un facteur de mérite M^* (7) par:

$$M^*_{\lambda_1, \lambda_2}(\Omega_v) = \int_{\lambda_1}^{\lambda_2} D^*_\lambda(\Omega_v) \left(\frac{\delta N(\lambda)}{\delta T} \right) d\lambda$$

$N(\lambda)$ étant la radiance spectrale du corps noir à la température To

$D^*_\lambda(\Omega_v)$ étant la détectivité du senseur

T étant la température

$M^*_{\lambda_1, \lambda_2}(\Omega_v)$ s'exprime en cm^{-1} Hz$^{1/2}$ sr^{-1} K^{-1}

Cette expression permet de calculer le contraste de température minimum détectable (ou équivalent au bruit) par:

$$CTEB = \frac{(A_d \, \Delta f)^{1/2}}{A_c \, \Omega_R \times M^*_{\lambda_1, \lambda_2}(\Omega_v)}$$

Avec:

    Ad surface sensible du détecteur

    Ac surface de collection de l'optique

    Ω_R angle solide de résolution du système

    Δf bande passante du système

L'intérêt de ce facteur de mérite réside dans le fait qu'il décrit simultanément la sensibilité du récepteur et l'amplitude du signal en fonction de la longueur d'onde.

IV.3 - Comparaison entre les deux bandes spectrales

Nous avons tout d'abors calculé le facteur de mérite d'un détecteur idéal dont la sensibilité est limitée par le fond ambiant en prenant deux cas extrêmes, large bande spectrale et angle de vue de 65°, puis bande spectrale étroite (0,3 um) et angle de vue de 15°. Ces deux courbes sont représentées figures 5 et 6. On observe que le facteur de mérite **M*** en large bande est supérieur d'un ordre de grandeur dans la gamme 10 μm, par rapport à la gamme 5 μm. En bande étroite la différence n'est plus que de la moitié d'un ordre de grandeur.

Nous avons ensuite calculé le temps d'intégration maximum du signal et la dégradation du facteur de mérite dans ces deux hypothèses. En comparant le temps d'intégration au temps de transfert (1 μs environ) nous déduisons la dimension de la matrice.

En large bande, la figure 7 montre que le temps d'intégration du signal dans la bande 5 μm est suffisant pour la lecture d'une petite matrice (une centaine d'éléments) et que le facteur de mérite est peu dégradé de sorte qu'il sera possible d'obtenir une détection limitée par le fond ambiant.

Par contre, dans la bande 8-14 μm les temps d'intégration sont nettement plus faibles et du même ordre de grandeur que le temps de transfert, ce qui rend impossible la lecture d'une matrice.

En bande étroite, la figure 8 montre que dans la bande 3-5 μm le temps d'intégration est tel que l'on pourra lire des matrices à très grand nombre d'éléments (10 000 points et plus), la sensibilité étant encore voisine de la limite théorique.

Dans la bande 8-14 μm, les facteurs de mérite sont très dégradés tandis que le temps d'intégration ne croît pas suffisamment pour permettre la lecture de matrices importantes (une centaine d'éléments au maximum). Pour accroître le temps d'intégration et améliorer la sensibilité il faudrait abaisser la température de fonctionnement aux environs de 50°K.

IV.4 - Mode de lecture CCD ou CID

Dans une matrice à lecture CCD (8) les charges sont transférées séquentiellement jusqu'à l'amplificateur de lecture situé au bord de la matrice (figure 9). Dans la structure CID (figure 10) les charges sont injectées à travers le substrat lorsque l'on sélectionne l'élément pour la lecture (9) (10). Un élément est constitué par deux électrodes, l'une connectée à une rangée d'adressage, l'autre à une colonne d'adressage. On obtient l'injection de cet élément en sélectionnant la ligne et la colonne. Lorsque l'une ou l'autre des deux électrodes est sélectionnée les charges sont transférées sous l'électrode non sélectionnée.

Les différences essentielles entre ces deux modes de lecture sont:

La lecture CID se satisfait d'une efficacité de transfert moins bonne que la lecture CCD, ceci étant particulièrement intéressant pour des matériaux semiconducteurs, tels que le tellurure de mercure cadmium dont la technologie de surface est moins bien connue que celle du silicium.

L'architecture d'une structure CID permet une excellente protection contre l'éblouissement dans les deux directions.

La technologie CID ne nécessite pas de diffusion ou d'implantation ce qui, compte tenu du fait que le nombre de couches n'est pas supérieur à celui d'un CCD, la rend plus facile.

Le rapport surface sensible sur surface totale de la matrice est plus important en CID qu'en CCD.

En contrepartie de ces avantages la structure CID nécessite l'emploi de multiplexeurs pour sélectionner les rangées et les colonnes. Ces circuits périphériques devant être placés au voisinage de la matrice pour limiter l'influence néfaste des capacités parasites. De plus la lecture CID ne se prête pas facilement au traitement par intégration avec délai, qui se réalise aisément dans le cas d'une lecture CCD.

V - CONCLUSIONS: DEFINITION D'UNE MATRICE 8 x 8

A partir des résultats de cette étude nous pouvons effectuer des choix pour la réalisation d'une matrice 8 x 8 éléments (Figure 11).

D'une part la lecture CID nous semble préférable à la lecture CCD. D'autre part compte tenu des temps de transfert et des temps d'intégration dans la gamme 3-5 um il est possible de réaliser une matrice 8 x 8 avec un seul circuit de lecture, en limitant le flux des photons pour obtenir un temps d'intégration de l'ordre de 500 μs. La longueur de l'électrode peut atteindre 50 μm. La largeur de l'électrode et l'espacement seront définis par la résolution optique désirée.

Dans la bande 8-14 μm, la matrice est décomposée en 4 matrices de 4 x 4 associées chacune à un circuit de lecture. On cherchera, par réduction du fond ambiant à obtenir un temps d'intégration de l'ordre d'une trentaine de microsecondes permettant de lire chacune des sous-matrices. La longueur des éléments étant de l'ordre de 30 μm pour que le transfert s'effectue en moins d'une microseconde. Les quatre circuits de lectures parallèles sont multiplexés à la sortie pour fournir le signal correspondant aux 64 éléments.

BIBLIOGRAPHIE

(1) S.M. SZE (1969) "Physics of Semiconductor Devices"
Wiley - Interscience

(2) FARRE J., BUXO J., ESTEVE D., C.R. Acad. Sc. PARIS (1976)
"Importance relative des composantes de courant de charge dans les capa-
cités MIS destinées à détecter le rayonnement infrarouge"
t. 283, Série B, 103-105

(3) ANDERSON W.W (1977) "Tunnel Current limitations of narrow bandgap infrared
charge coupled devices"
Infrared Physics, Vol.17, 147-164

(4) BAYLE, W.S and SMITH G.E.(1970) "Charge Coupled Semiconductor Devices"
B.S.T.J. 49, 587-593

(5) BAYLE, W.S. and SMITH G.E.(1971) "Charge Coupled Devices - A new approach to
MIS Devices Structures"
IEEE Spectrum 8, 18-27

(6) NGUYEN-DUY, M. ROYER - Communication privée

(7) J.W. SABEY (1975) "Infrared detector figures of merit for broadband sensing
applications"
I.E.E International Conference on Low light and thermal imaging systems
LONDON

(8) For a review, SEQUIN C.M. and TOMPSETT M.F. (1975) "Charge transfer devices"
Academic Press. Inc.

(9) MICHON G.J. and BURKE H.K (1973) "Charge injection imaging"
Dig. Techn. Papers Int. Solid - State Circuits Conf. - 138-139

(10) MICHON G.J. and BURKE H.K. (1974) "Operational Characteristics of CID imager"
Dig. Techn. Papers Int. Solid - State Circuits Conf. - 26-27

DISCUSSION

Steckel

Have you characterized the surface recombination velocity for any of the insulators you mentioned.

Réponse d'Auteur

Nous n'avons pas actuellement fait de mesures très précises sur les vitesses de recombinaison en surface de ces dispositifs. Les seules mesures sont des mesures indirectes à partir de celles que nous faisons sur le temps de relaxation de la structure à l'obscurité. Nous déduisons la vitesse de recombinaison en surface de la mesure du temps de relaxation de la caractéristique Capacité en fonction du temps.

Flux de photons V_G

Métal

Isolant

Couche d'inversion

J_P J_G

Zone dépeuplée

J_D

Semiconducteur Type N
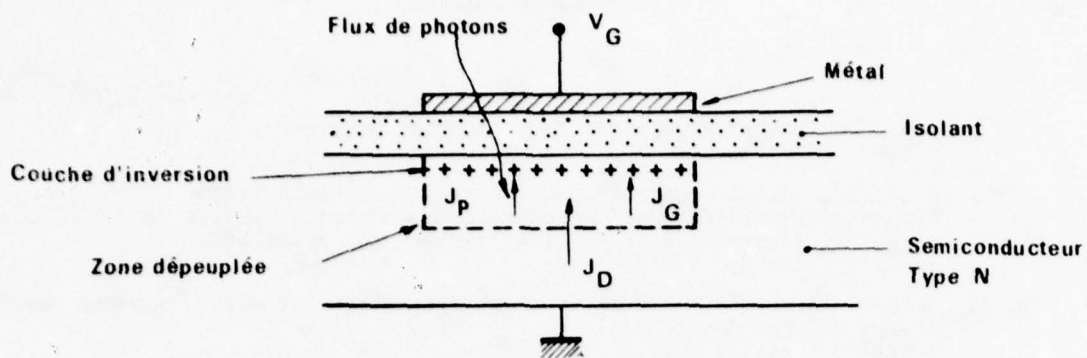
(a) Coupe d'une structure M. I. S.

J_T : courant tunnel

J_G : courant de génération thermique

J_D : courant de diffusion

J_P : courant de génération photonique

Isolant

Métal

Semiconducteur type N

E_{FM}

J_T

$J_G / 2$

$J_P / 2$

$q \, \phi_S^0$

Couche d'inversion

Zone dépeuplée

Photons

E_C

E_F

Eg

$J_G / 2$

$J_P / 2$

J_D

E_V

W
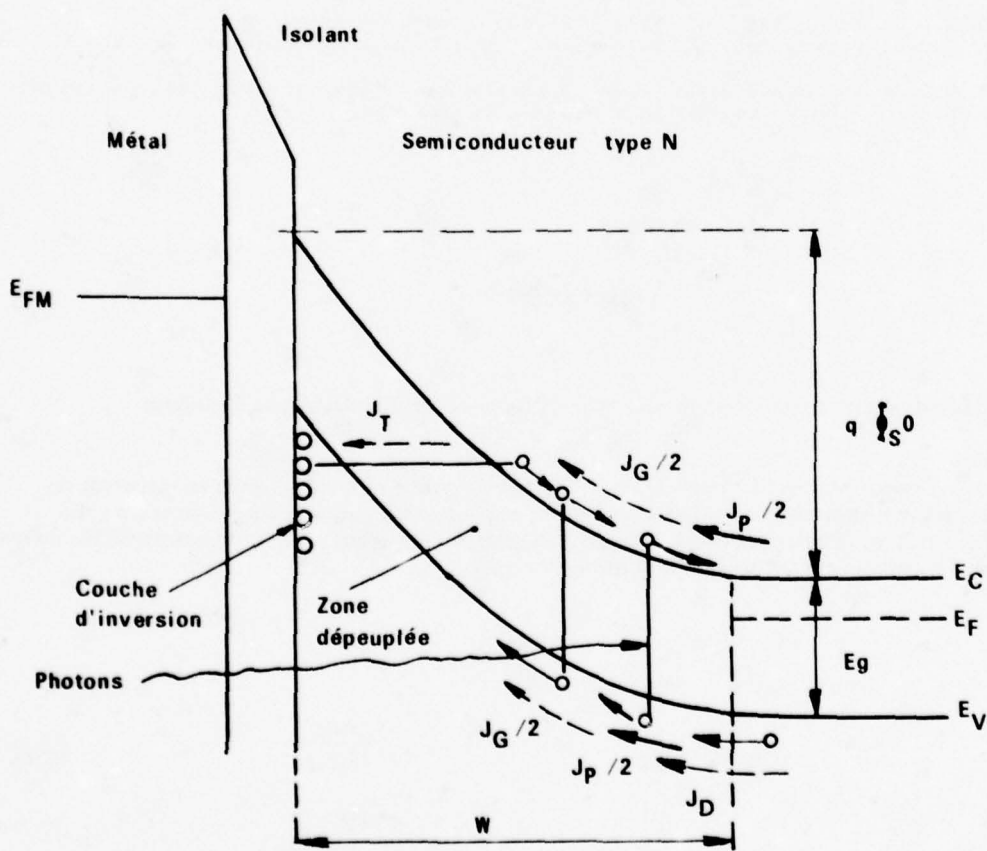
(a) Diagramme d'énergie

Figure 1   : Principe d'une structure M.I.S.

Temps de relaxation thermique en s.

Température : 77°K

Dopage : N_D 10^{15} cm^{-3}

Durée de vie τ : 10^{-6} s

Isolant : (SiO) 1000 Å

S : vitesse de recombinaison en surface

S = 200 cm/s

S = 1000 cm/s

Longueur d'onde de coupure μm
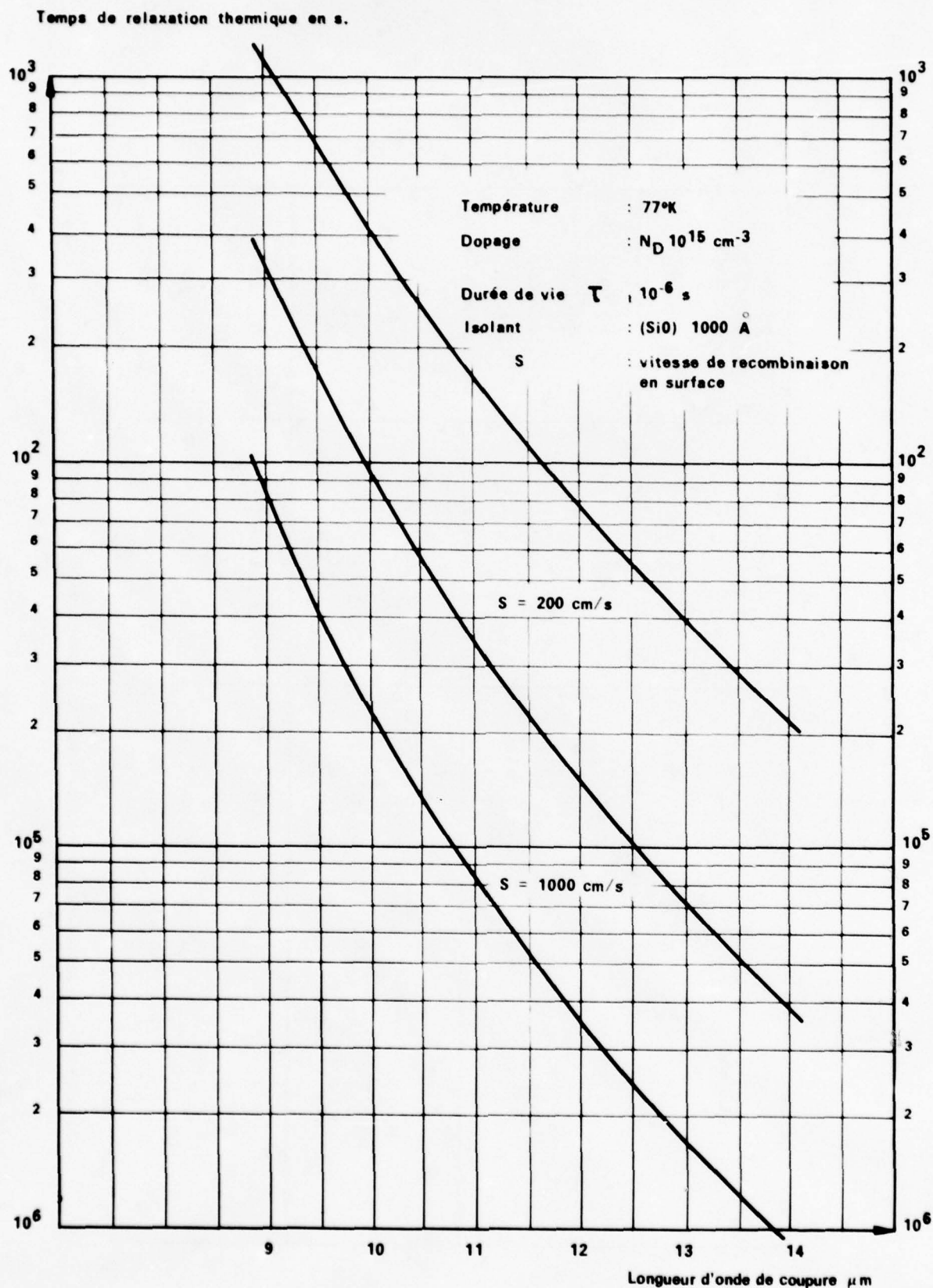
Figure 2 : Variation du temps de relaxation thermique avec la longueur d'onde de coupure et la vitesse de recombinaison en surface

LONGUEUR D'ONDE : 10 μm

Isolant	Dopage N_D (cm^{-3})	Charge maximale N_{max} (cm^{-3})	Temps de relaxation thermique T_R (S=1000)
SiO $\varepsilon = 5,5$	2.10^{15}	4.10^{10}	10 μs
	10^{15}	10^{11}	24 μs
	5.10^{14}	2.10^{11}	40 μs
	2.10^{14}	5.10^{11}	85 μs
Oxyde anodique $\varepsilon = 11$	2.10^{15}	5.10^{10}	17 μs
	10^{15}	$1,3.10^{11}$	40 μs
	5.10^{14}	3.10^{11}	75 μs
	2.10^{14}	9.10^{11}	160 μs
TiO$_2$ $\varepsilon = 75$	2.10^{15}	$1,3.10^{11}$	83 μs
	10^{15}	6.10^{11}	200 μs
	5.10^{14}	$1,7.10^{12}$	450 μs
	2.10^{14}	6.10^{12}	1 ms

Figure 3 : Influence du dopage et de la permittivité de l'isolant sur le temps de relaxation thermique de la structure M.I.S Cd Hg Te
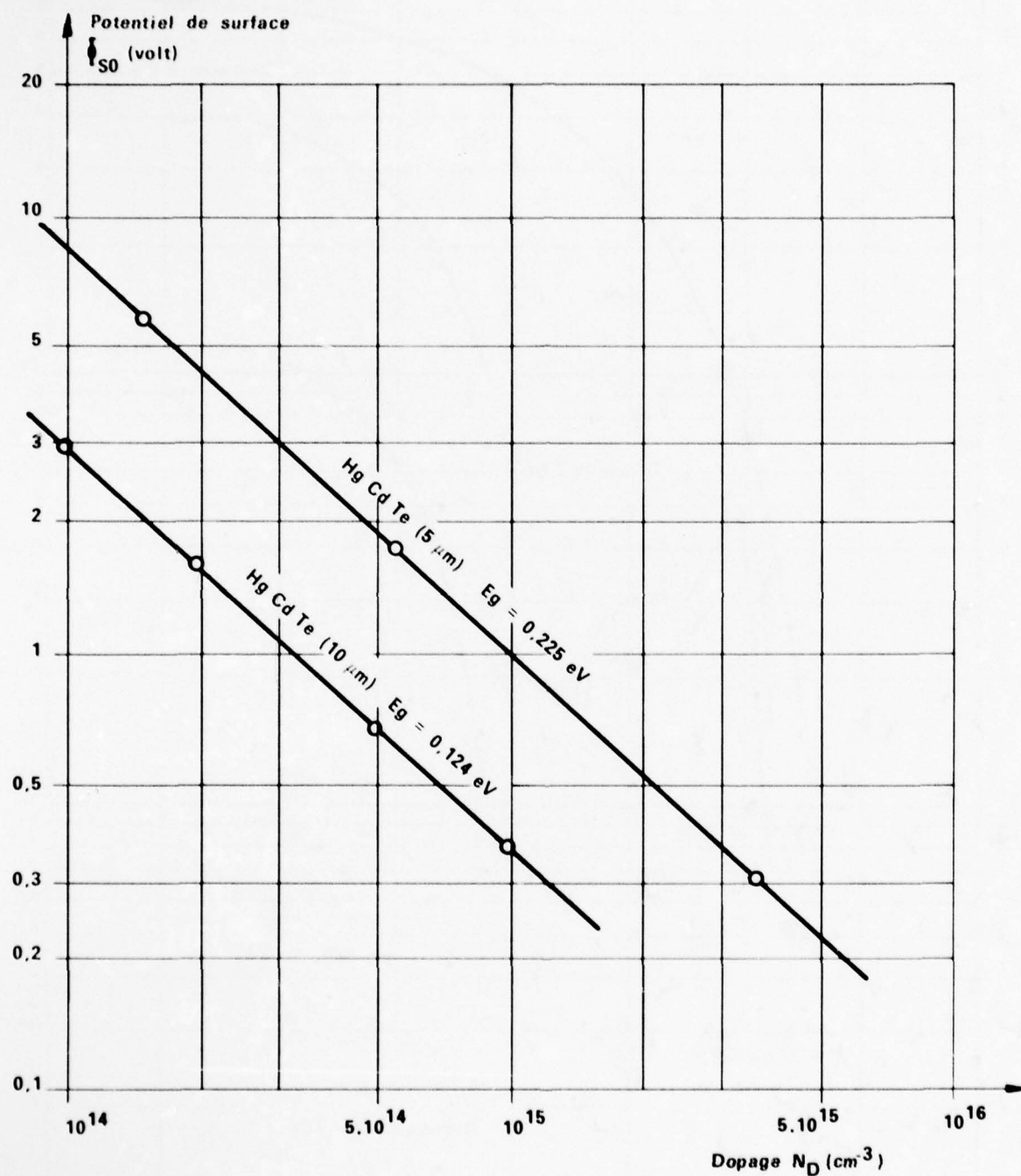
Figure 4 : Potentiel de surface maximum compte tenu de la limitation par l'effet tunnel en fonction du dopage
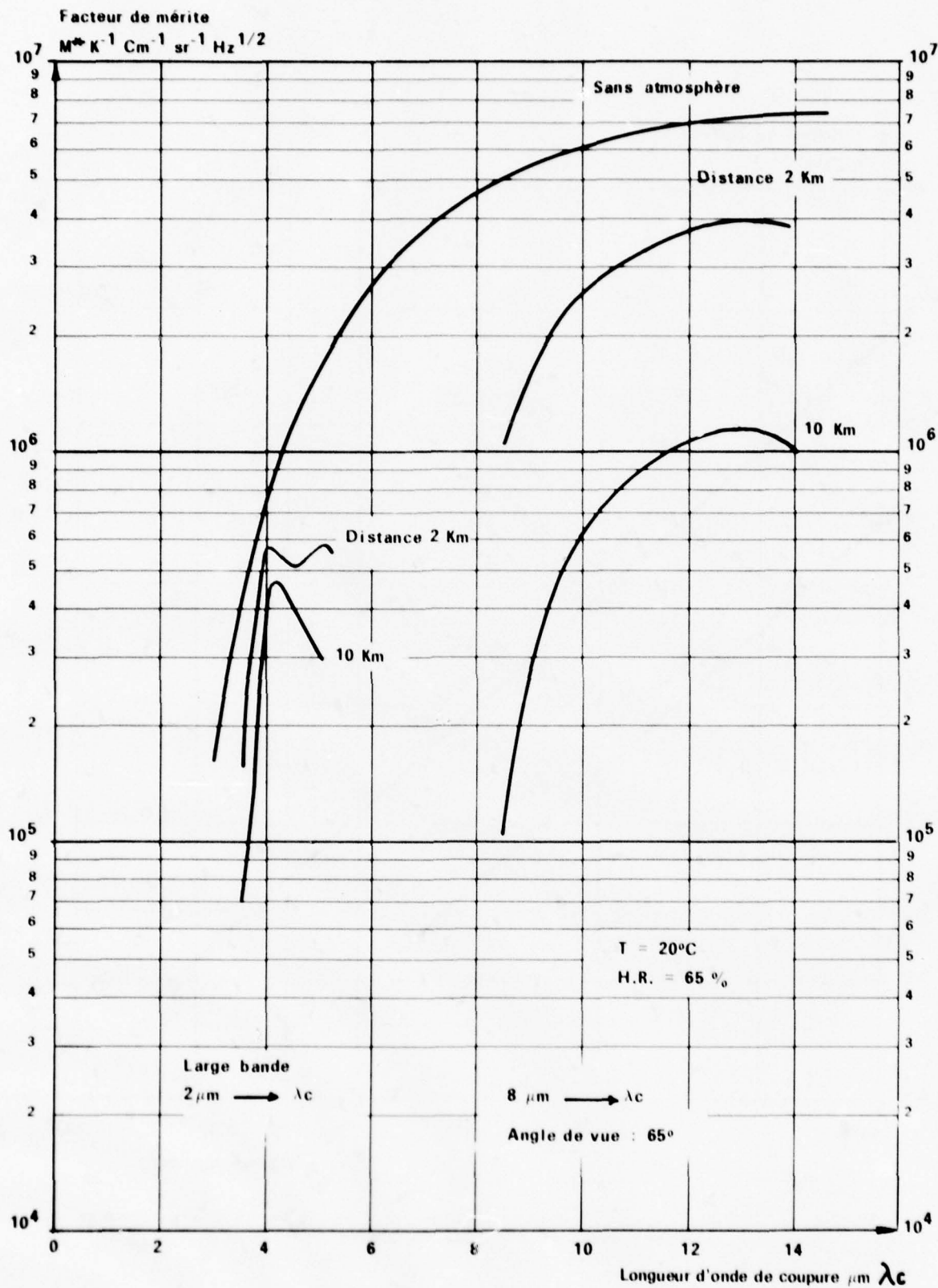
Figure 5 : Variation du facteur de mérite en fonction de la longueur d'onde de coupure et de l'absorption atmosphérique en bande large

Facteur de mérite

$M^X K^{-1} Cm^{-1} sr^{-1} Hz^{1/2}$

Sans atmosphère

R = 2 Km

R = 10 Km

R = 2 Km

R = 10 Km

T = 20ºC

H.R. = 65 %

Bande étroite
$\Delta \lambda = 0.3 \mu m$

Angle de vue : 15º

Longueur d'onde de pic en μm
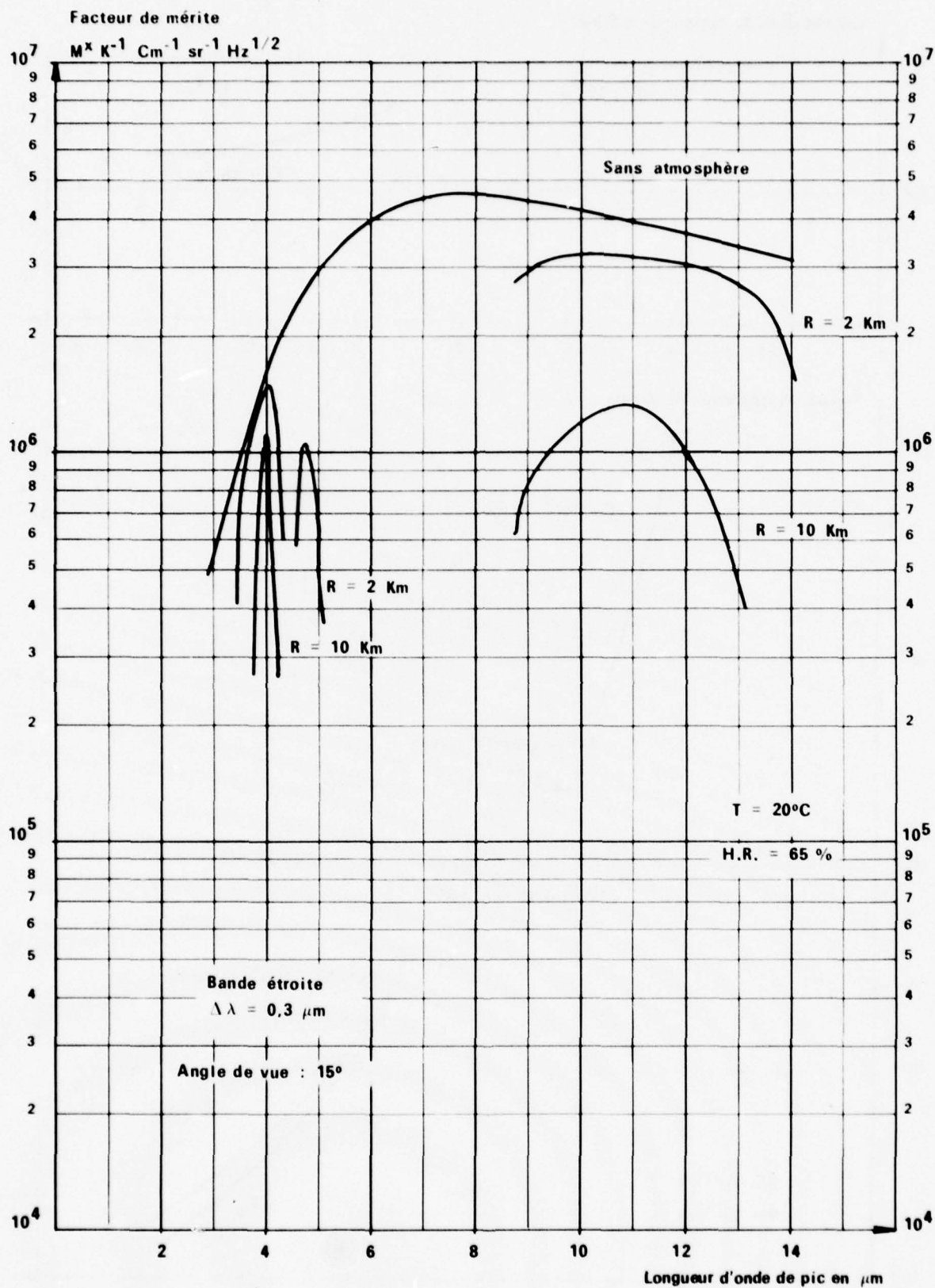
Figure 6 : Variation du facteur de mérite en fonction de la longueur d'onde de coupure et de l'absorption atmosphérique en bande étroite
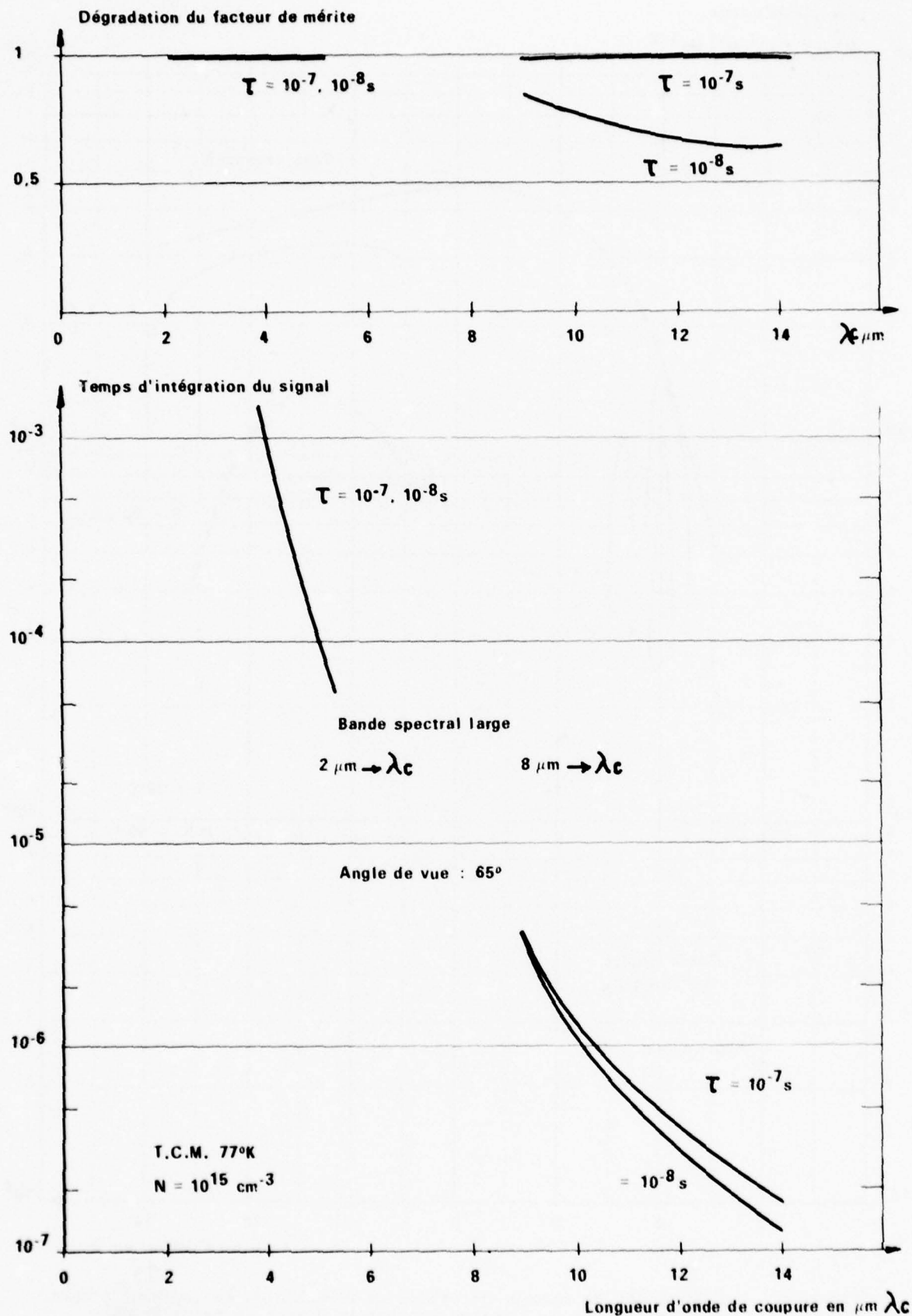
Figure 7 : Résultats de la modélisation en bande spectrale large
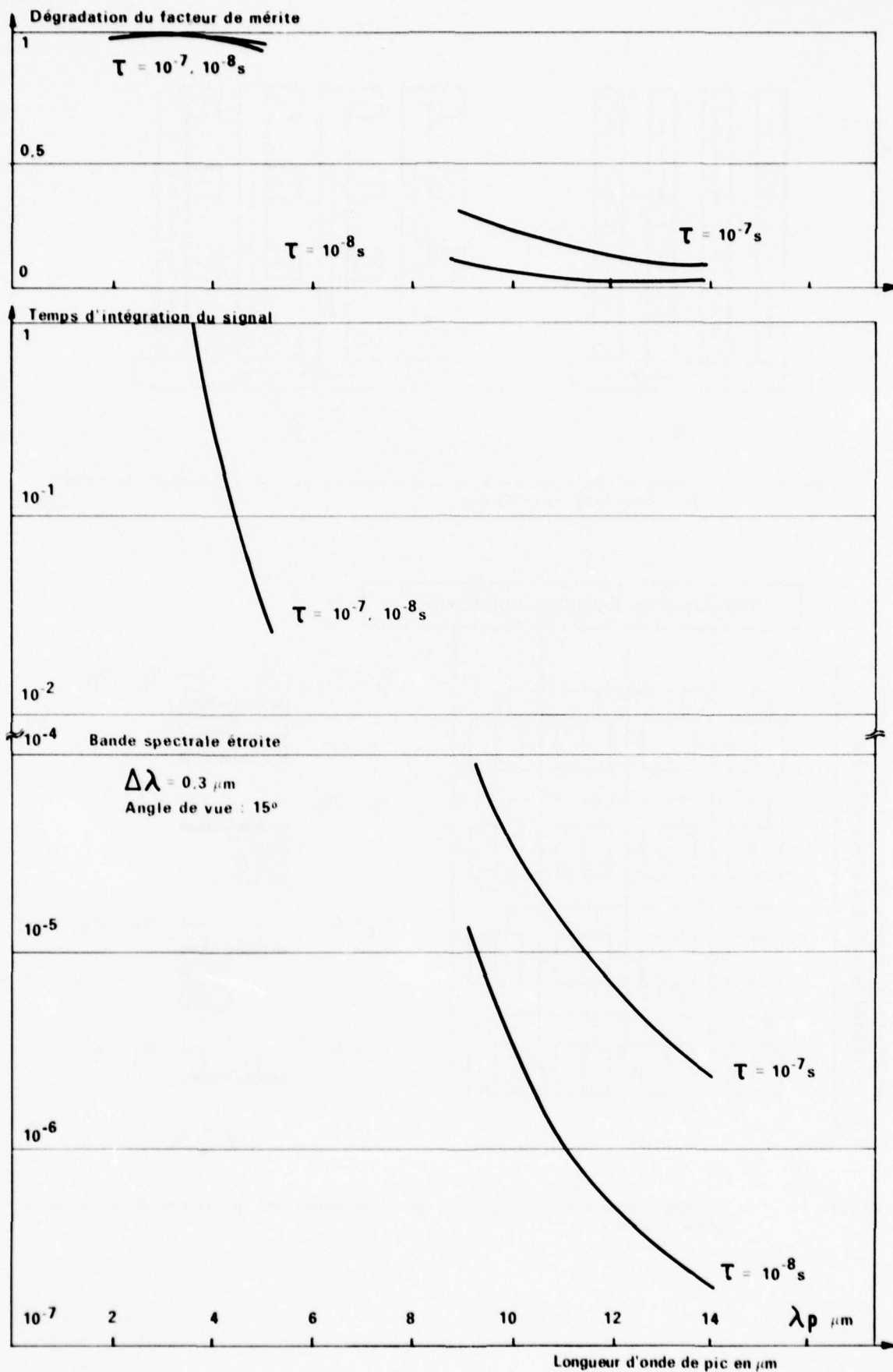
Figure 8 : Résultats de la modélisation en bande spectrale étroite

(a)

(b)
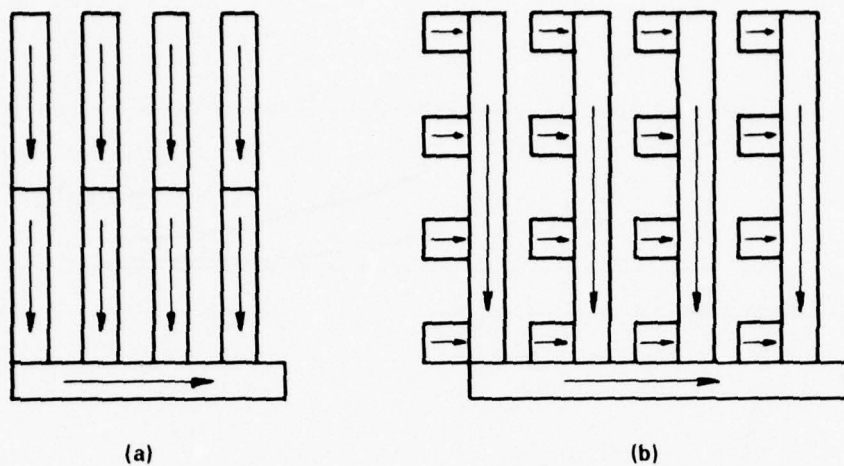
Figure 9 : Schéma d'une matrice à architecture CCD: (a) transfert de trame, (b) transfert interligne

GENERATION DU BALAYAGE HORIZONTAL

GENERATION DU BALAYAGE VERTICAL

$V_X = V_D$ $V_Y = V_D$ 1

$V_X = V_D$ $V_Y = 0$ 2

$V_X = 0$ $V_Y = V_D$ 3

$V_X = 0$ $V_Y = 0$ 4

(d)
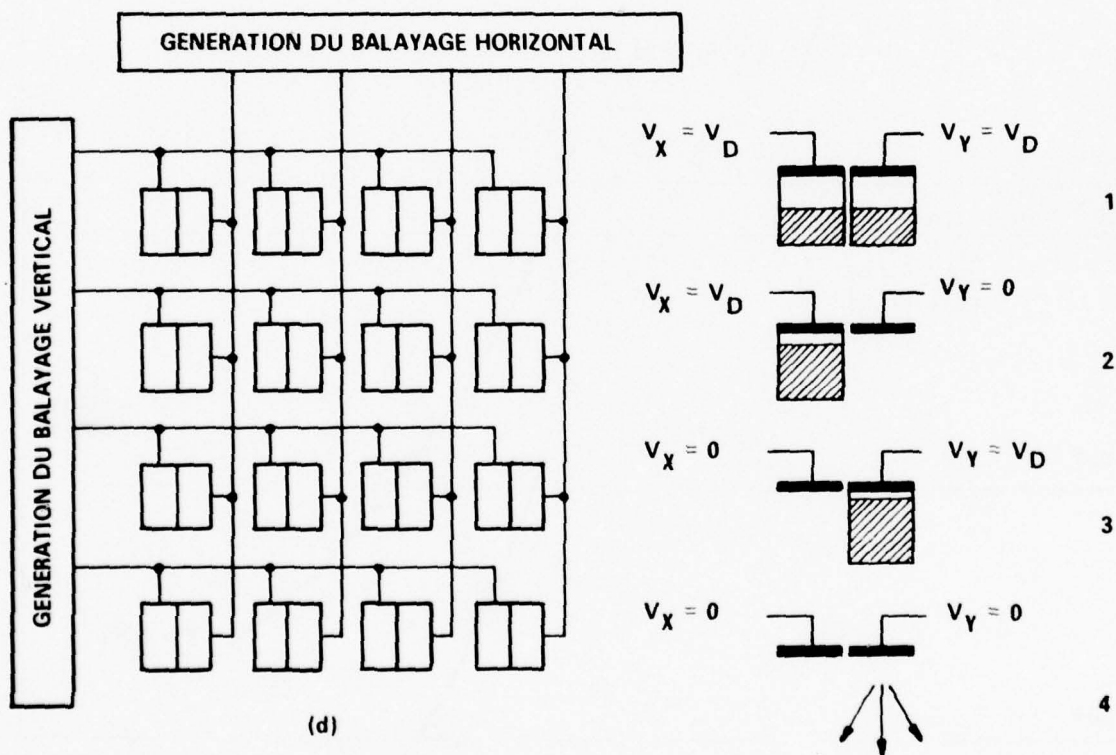
Figure 10 : Schéma d'une matrice CID (a) architecture, (b) principe de sélection pour lecture
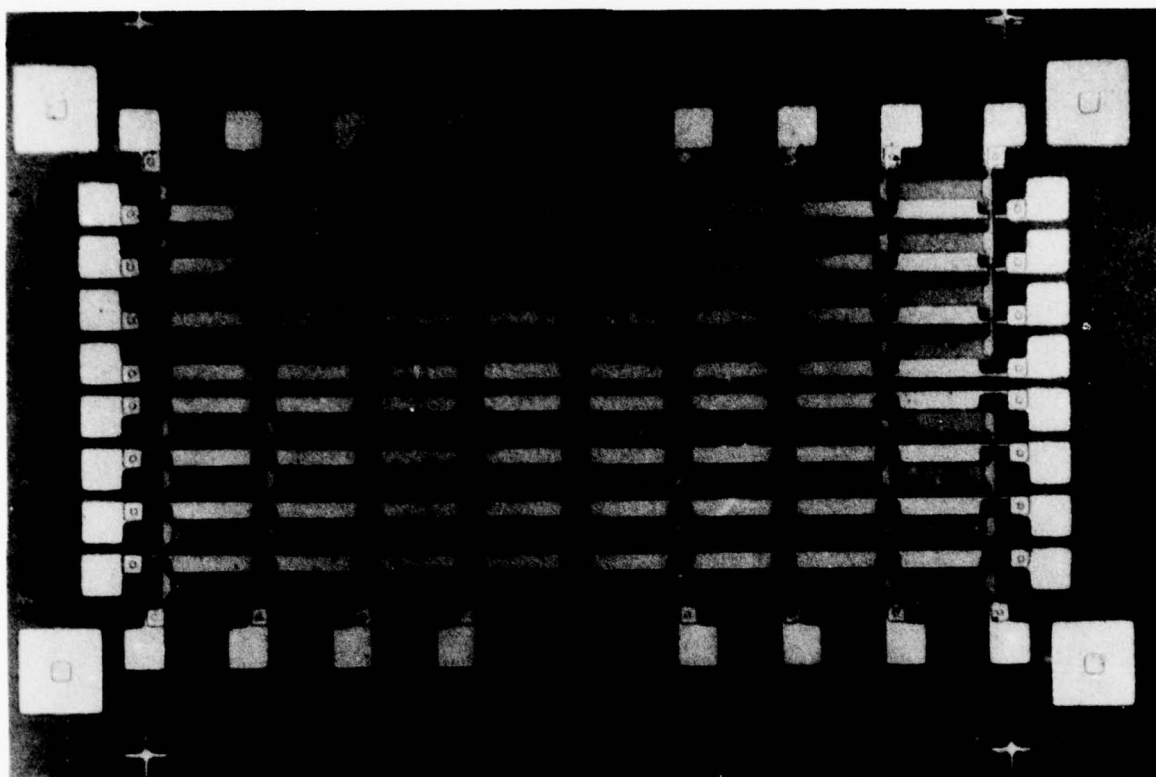
Figure 11 : Photographie d'une matrice 8 x 8 à lecture CID.

IRCCD IMAGING SENSORS: A REVIEW OF DEVICE OPTIONS

Andrew J. Steckl
Rensselaer Polytechnic Institute
Troy, New York 12181
USA

ABSTRACT

In this paper a review is presented of the many device options presently being investigated for the development of integrated infrared focal plane arrays. With only one or two exceptions, the approaches being developed for this goal are centered on the charge coupled device (CCD) concept or its close relative, the charge injection device (CID). Monolithic and hybrid IRCCD and IRCID device options are classified, discussed and compared.

1. INTRODUCTION

The invention of charge-coupled devices (CCDs)(BOYLE, 1970) has brought a new perspective to the development of infrared optical imaging and detection systems. It impacts both the need for ever increasing detectivity and spatial resolution, and the wider use and deployment of such systems through lower unit cost, component miniaturization, and relaxed operating requirements.

The charge-coupling concept is based upon the formation and transfer of charge packets in a repetitive metal-insulator-semiconductor (MIS) structure. When charge packets are photogenerated in the substrate and then electronically manipulated to a single output, the CCD operates as the solid-state, self-scanned equivalent of a vidicon tube. When the charge packets are introduced electronically, the CCD operates as a sequential shift register in which the time delay can be electronically varied by changing the clock frequency. The simplicity of the basic CCD combined with the fact that it can be operated as an imager and/or shift register has resulted in a remarkably versatile device.

2. IRCCD OPERATION AND DEVICE OPTIONS

The application of CCDs to the development of a new generation of infrared imaging sensors is receiving considerable support and has met with significant successes. In previous IRCCD reviews (STECKL, 1975a, 1976) undertaken when the field was quite young, we suggested that the many different device approaches to the goal of integrating IR detector arrays with CCDs can be grouped into two major classes: monolithic and hybrid devices. Today, evaluating a more mature field we feel that this division is still essentially correct.

In general, monolithic IRCCDs have in common the central premise that the closest and most efficient integration would be achieved by fabricating both the photosensitive section (i.e., the IR array) and the read-out section within one structure. By contrast, the hybrid IRCCD category consists of devices in which the major functions occur in distinct but integratable components and materials. Photodetection and charge generation in a hybrid device can take place in any one of various suitable IR photodetectors. The signal read-out takes place in a standard silicon CCD or possibly an FET or diode X-Y address matrix.

To the monolithic and hybrid IRCCDs one must now add a third category, namely the infrared charge injection device (IRCID). In the CID, the photogeneration and part of the charge handling functions are performed in a monolithic array, while the actual read-out is performed by one or two silicon shift registers. The CID approach is further discussed in Section 5.

Figure 1 covers the various device options which fall in the three major categories outlined above. In the monolithic IRCCD category are included:

(a) A narrow band gap semiconductor substrate and conventional CCD structure.
(b) A narrow band gap semiconductor-silicon heterojunction and a conventional CCD structure fabricated on the Si substrate.
(c) A metal-silicon Schottky barrier structure with a conventional CCD structure fabricated on the Si substrate.
(d) An extrinsic Si substrate on which an accumulation mode CCD is fabricated.
(e) An extrinsic Si substrate on which an epitaxial layer is grown and on which a conventional CCD is fabricated.

In the hybrid IRCCD category are those devices which combine a Si CCD read-out with the following types of detectors:

(a) Photoconductive detectors.
(b) Photovoltaic detectors.
(c) Pyroelectric detectors.

Finally in the last category fall those devices which use the following substrates for the fabrication of a Charge Injection Device Array:

(a) A narrow band gap semiconductor substrate.
(b) An extrinsic silicon substrate.

A major difference between the operation of CCDs in the visible and in the infrared is the large background flux rate Q_B in the infrared (Fig. 2). This results in increasingly shorter background charge-up or saturation time with increasing wavelength of operation (from about 1 sec for the 2.0 to 2.5 μm window to about 10 μsec for the 8 to 14 μm region). Thus, one limitation of practical IRCCD devices is the length of exposure time.

Another important consideration, especially for IR thermal imaging applications, is the small contrast between the photon flux generated by an incremental change in background temperature and the total photon flux. For 0.1°K change (typical for IR imaging) of the average 300°K background, the contrast is on the order of 1.0% in the 2.0 to 2.5 μm window and less than 0.1% in the 8 to 14 μm region (Fig. 3). The presence of such low contrasts points to another practical limitation: the uniformity of IR material properties. Nonuniformity in the responsivity of IR detectors, of the same order or higher than the scene contrast, seriously degrades the minimum resolvable temperature and results in severe fixed pattern noise.

A number of IRCCD focal plane organizations have been explored (Fig. 4). In the serial scan approach, a linear array of detectors is raster scanned across the scene, the direction being parallel to the array (Fig. 4a). By processing the outputs of the detectors through a time-delay-and-integration (TDI) operation, the components of the signals corresponding to the same resolution element are summed linearly.

With a CCD, TDI operation is achieved by clocking the device at a rate corresponding to the image velocity across the detector array (Fig. 5). Because the noise contribution of each detector is independent and thus uncorrelated, the total noise is obtained by an rms summation. For an array of m detectors, the TDI operation could, therefore, result in a potential maximum improvement of $\sqrt{m}$ (in the detector-noise-limited case) in the signal-to-noise ragio (SNR) of the entire array over that of an individual detector. The upper limit of m and, therefore, of the maximum achievable SNR improvement is set by the total detector array signal versus the CCD charge capacity. The operation of the entire array appears at the CCD output essentially as one detector. This results in reduced fixed-pattern noise, lowered detector response uniformity requirements, and built-in redundancy. These features, together with the potential $\sqrt{m}$ improvement in SNR, help circumvent the problems posed by the high background and low contrast present in thermal imaging systems at IR wavelengths.

In the parallel scan FLIR a vertical linear array of detectors is slowly scanned across the scene, the scan direction being perpendicular to the array (Fig. 4b). In this case, an input-tapped CCD shift-register is used as a multiplexer which reads out the information line by line. The serial-parallel FLIR consists of a two-dimensional array which is scanned across the scene. TDI is performed in each channel along the scan direction and all channels are read out by a multiplexer. The "staring" two-dimensional array fully covering the entire scene field of view (Fig. 4c) requires no scanning and is the eventual goal of IRCCD focal planes.

3. MONOLITHIC IRCCDs

The various monolithic IRCCD options have in common the basic feature of combining both the detection and read-out in one structure. On the other hand, the feature which distinguishes one monolithic approach from another is the basic photodetection mechanism.

The two basic carrier photogeneration mechanisms are intrinsic generation and impurity ionization. These two mechanisms are illustrated in Fig. 6. In the intrinsic absorption case, photons with energy equal to or greater than the band gap are absorbed and an electron-hole pair is generated. In the extrinsic absorption case, photons with energy equal to or greater than the ionization energy of a given impurity are absorbed resulting in one mobile carrier and a fixed ion. In both cases. the photocarriers are collected by a built-in or applied electric field.

Investigations of monolithic narrow gap semiconductor IRCCDs have been reported for a broad range of materials, Fig. 7. For 3-5 μm wavelength operation the binary compounds PbS, PbSe, PbTe and InSb are presently under investigation. In addition, the ternary alloys InAsSb and HgCdTe are also being considered due to their ability to closely match a desired spectral response. In the far infrared, at 8-12 μm, ternary alloys of HgCdTe and PbSnTe can also be considered. However, due to the material problems (e.g., dark current, nonuniformity, etc.) the present emphasis is on the 3-5 μm band.

The choice of developing a particular monolithic narrow-gap semiconductor IRCCD basically depends on the state-of-art of the particular semiconductor. However, it is important to highlight which material properties of the semiconductor affect various device characteristics. In Fig. 8, the interaction between material and operation parameters and device charcteristics is elaborated. The five basic device performance criteria chosen are:

(1) storage time
(2) transfer time
(3) quantum efficiency
(4) dynamic range
(5) cross-talk

A sixth important device criterion, noise, was not included since it depends as much on device design as on material characteristics.

The device storage time (STECKL, 1975a) is probably the characteristic which is dependent on most material and operational parameters, most prominently on the minority carrier lifetime, the carrier concentration, the insulator dielectric constant, the device temperature, the photon flux and the gate voltage. Similarly, the transfer time or the transfer efficiency of the device is dependent on a sizable number of material and operational parameters: the inversion-region mobility, the surface-state density, the temperature, the gate voltage, the gate length, etc. On the other hand, the quantum efficiency of the device is basically only a function of the semiconductor absorption coefficient and the volume from which photogenerated carriers can be collected. The dynamic range of the device is given by the ratio of the maximum number of carriers which can be stored in a cell to the number of noise carriers in that cell. The upper limit of the dynamic range is a function of the capacity of the cell and, therefore, depends on the dielectric constant of the insulator and the applied gate voltage. Finally, the last device parameter, the cross-talk, is somewhat more complex in its dependence on various parameters than Fig. 8 indicates. However, optical cross-talk is

basically a function of the number of photocarriers collected in pixels adjacent to the one on which the radiation is incident. Therefore, cross-talk is very dependent on the absorption coefficient and on the size of the pixel.

A major difficulty in implementing the monolithic narrow-gap semiconductor IRCCD is the passivation of the semiconductor-insulator interface. As opposed to silicon, most of the narrow-gap materials do not possess a native oxide which could result in a low number of surface states. However, this technology is presently being developed for selected semiconductors and it is probable that this obstacle can be overcome. As an example, well-behaved MIS structures have been fabricated (KIM, 1975) on InSb using SiON as the insulator. The capacitance-voltage characteristics of an InSb MIS device are shown in Fig. 9 and it can be seen that they follow classical behavior.

The monolithic extrinsic semiconductor IRCCD development has to date been concentrated on doped silicon in order to take full advantage of its well developed technology and associated possibilities of low-cost batch processing. Because of the maturity of IC technology, the emphasis in Si:X IRCCD work is on the development of large-scale arrays using such sophisticated tools as computer-aided design and automated processing and testing. At the same time, by the use of batch processing one can expect that the eventual unit cost of extrinsic silicon detectors will be substantially reduced.

The main drawback of the Si:X IRCCD is the low temperature of operation required by Si:X detectors. For the 8 to 14 μm region operation, a temperature between 15 and 30°K is generally required, while for the 3 to 5 μm region the temperature range is typically between 40 and 65°K. This represents a serious problem in certain applications and higher cooling system cost in most applications. The Si:X IRCCD also inherits the relatively low quantum efficiency ($\sim$5 to 30%) associated with the long impurity absorption length in extrinsic detectors. The effect in turn, can result in increased cross-talk due to repeated reflection of radiation within a detector array.

In Fig. 10 are shown the impurity ionization energies and corresponding cut-off wavelengths of extrinsic silicon and germanium photoconductors. The common Si impurities vary in ionization energy from around 45 meV for P and B to approximately 150 meV for In. As can be seen the match between the cut-off wavelength of these impurities and either the 3-5 μm or 8-14 μm window is not as good as it is in the case of certain narrow band gap semiconductors. For that reason, other, more exotic impurities are presently investigated.

Since extrinsic detectors are by their very nature majority carrier devices, and CCDs use minority transport, the normal CCD structure is not adequate for a Si:X IRCCD. A number of devices have been conceived and are being studied which resolve this incompatibility.

The majority carrier or accumulation mode CCD (AMCCD) is one approach. Accumulation-mode operation of an MIS structure results when the proper polarity gate voltage induces the same majority carrier type as in the bulk of the semiconductor. The AMCCD operation requires operation at a temperature low enough such that thermal generation is negligible. The transfer efficiency of AMCCDs (NELSON, 1974) has been found to be around 0.9 at best which is two orders of magnitude less efficient than inversion-mode devices operated (STECKL, 1975b) at the same temperature.

A second approach is to physically separate the photosensitive area from the CCD register (Fig. 11) and to manipulate the signal such that the right polarity charge is introduced into the CCD. This horizontally integrated approach has substantial dead space, which is a drawback for two dimensional arrays but not for line imagers.

A third approach is to vertically separate the photodetection and read-out functions. This can be achieved by growing an epitaxial layer of opposite type to the substrate and then fabricate the CCD on this layer (Fig. 12). In this fashion, when photogenerated majority carriers from the substrate are injected into the epilayer, (Fig. 12b), they represent minority carriers and can be handled in the usual fashion. Another solution is to invert the photocurrent polarity before introducing it into the CCD proper (Fig. 12c). This buffer circuit can be designed and processed simultaneously with the CCD, thus retaining the total integrated-circuit aspect of the Si:X IRCCD. The special impurities used in Si:X detectors (e.g., Ga, Al, and In) require special processing techniques when combined with the usual Si IC impurities (e.g., P, As, and B). Furthermore, the special device configurations, such as the growth of an epilayer and subsequent CCD fabrication, represent additional processing compatibility requirements.

Another monolithic IRCCD approach is based on internal photoemission from metal-semiconductor Schottky barrier arrays on a silicon substrate (KOHN, 1976). The photoemission process is shown in Fig. 13. Photons with energy $h\nu < E_g$ are absorbed in the metal resulting in the excitation of hot carriers. Carriers with energy larger than the contact barrier, and with sufficient momentum in that direction, are emitted into the semiconductor. With radiation incident on the Schottky barrier through the semiconductor, the photogeneration process depends only on the absorption in the metal and emission over the barrier, resulting in uniform responsivity. The quantum efficiency of the Schottky barrier detector decreases with decreasing photon energy and is quite small at the IR wavelengths of interest. For example, for an Au-p-Si diode, a quantum efficiency of 1% at 3 μm was reported (KOHN, 1976). However, the combined characteristics of high uniformity and low responsivity can be potentially useful in large staring arrays where the longer integration time increases the detectivity.

The last monolithic IRCCD concept to be discussed is centered around a heterojunction structure consisting of a thin narrow-gap semiconductor film on a Si substrate (STECKL, 1977). This approach is particularly suitable to a number of lead salt semiconductors (e.g., PbS, PbSe, etc.) which can easily be deposited on the Si substrate. At IR wavelengths photogenerated carriers in the lead salt film are injected into the Si (Fig. 14) where they are first collected and then read-out by a CCD structure.

4. HYBRID IRCCDs

Since the hybrid IRCCD involves basically the coupling of two fairly well developed technologies, it initially received more attention than monolithic IRCCDs.

Initial work on hybrid devices centered on the serially-scanned FLIR where it was quickly realized (STECKL, 1973, Erb, 1973) that the large IR background and small temperature contrast do not represent insurmountable difficulties. The serially-scanned FLIR uses a linear array of detectors raster-scanned across the scene, the scan direction being parallel to the array. Since each detector in turn scans the entire field of view, the dwell time per resolution element is quite short (Fig. 15). The amount of charge generated during this dwell time in even an 8 to 14 μm IR detector by a 300°K blackbody is well within the charge handling capability of a typical CCD shift register.

The two basic methods for introducing IR signals from the detector array into the tapped CCD shift register are the direct coupled input and the buffer interface input (Fig. 16). In the direct coupled input, the detector is connected either to the input diffusion (direct injection) or to the input gate (gate modulation) of the CCD tap. In the second method, the IR signal is first processed through a buffer stage before being introduced into the CCD.

In the direct injection IRCCD (STECKL, 1973) the photogenerated charge is directly introduced into the CCD shift register. Since this is in effect a dc coupled system, only IR detectors which exhibit relatively small dc currents (e.g., photovoltaic, extrinsic detectors) can be coupled to the CCD due to the latter's limited charge handling capacity.

Figure 17 illustrates the basic injection concept for a hybrid IRCCD consisting of a (Pb,Sn)Te photodiode and an n-channel CCD. The (Pb,Sn)Te/PbTe heterostructure is particularly attractive for a hybrid IRCCD array since integration can be achieved in a relatively simple sandwich structure with full use of the detector active area and requiring no interconnects. As shown in the figure, the IR diode is connected in parallel to a silicon diode which serves as an input tap to the CCD. The first MOS gate, V_G, serves to reverse bias both diodes. While the charge accumulates under the storage gate V_S, it is isolated from the CCD channel by the transfer gate V_T. After one read time, t_R, V_T is biased into inversion and the accumulated charge is transferred into the CCD channel. A critical parameter of the direct injection IRCCD is the injection efficiency, defined as the ratio of the charge introduced into the CCD to the total charge generated by the detector.

Hybrid IRCCDs using pyroelectric detectors are also being investigated. Pyroelectric materials owe their photodetection property to a temperature dependent polarization. In the steady state, the polarization is masked by surface charges and, therefore, the incident radiation must be chopped. This results in ac coupled detection and thus background subtraction. Additional features of pyroelectric detectors are inexpensive detector materials and room temperature operation. However, the sensitivity of present pyroelectric detectors is considerably lower than that of quantum detectors. Detection of the thermally induced electric polarization of the pyroelectric is accomplished by making the pyroelectric the dielectric of a capacitor. The two methods which we previously proposed (STECKL, 1975a) for introducing the pyroelectric signal into a CCD are probably still the most attractive: 1) connecting the capacitor to an on-chip MOSFET or 2) fabricating a pyroelectric film between the MOSFET channel and the gate metal (viz., in series with the gate). Both methods are essentially the same, the latter requiring more sophisticated processing, but resulting in a more compact structure.

5. CHARGE INJECTION DEVICES

The last major IRCCD category is based on the charge injection device (BURKE, 1976) (CID) concept. The CID unit cell consists of two overlapping electrodes. As in the CCD photogenerated minority carriers are collected at the semiconductor-insulator interface in a potential well underneath the electrodes. As one electrode in the cell is turned off the stored charges accumulate under the second electrode. When both electrodes are simultaneously switched off, the stored charge in the cell will be injected into the substrate. The collected charge can be sensed by integrating the substrate current or by monitoring the voltage on the electrodes.

A CID imager takes the form of an X-Y matrix array, where each pixel is sequentially addressed by a combination of two scan registers (Fig. 18). The charge in the pixels in the row and column not simultaneously addressed by both registers is simply transferred back and forth between the two electrodes of each pixel. Herein lies the advantage of the CID approach: since the charge is simply moved back and forth (rather than sequentially through the array as it is in the CCD) the transfer efficiency and, therefore, the surface state density does not play a major role. A CID approach is, therefore, very appealing for those materials which have surface state density too large for CCD operation. A second attraction of the CID approach is the fact that the pixels can be randomly as well as sequentially addressed.

It should also, however, be pointed out that a CID array is not self-scanned as an equivalent CCD array would be. Separate shift registers are required to scan the CID array, resulting in a mechanically hybrid focal plane.

The CID approach has been very successfully applied (KIM, 1977) to InSb. Arrays of 12 x 24 elements with 2 x 2 mils^2 cells have been fabricated and operated, which is indicative of the great potential of this approach.

6. SUMMARY AND CONCLUSIONS

In this paper a review of the various IRCCD device options has been presented. The emphasis in this review has been on present or future device techniques which could lead to the large scale integration of infrared sensors and not on the performance of present IRCCDs. Consequently, only a brief and qualitative summary of IRCCD work in progress is presented in Figs. 19 and 20 for monolithic and hybrid devices respectively.

In the monolithic narrow gap semiconductor IRCCD category, the development of InSb devices is furthest along. MIS, MOSFET and CCD operation have been demonstrated. Other semiconductors presently investigated include PbS, PbTe and HgCdTe.

The extrinsic Si monolithic IRCCD is probably the device option receiving the most attention. IRCCD arrays have been fabricated and operated using Si:Ga and Si:Al for the 8-12 μm and Si:In for the 3-5 μm regions.

Schottky barrier devices have been incorporated into an integrated IRCCD and operational feasibility has been demonstrated. In the heterojunction IRCCD, PbS-Si heterojunction detectors have been demonstrated and work on the development of an integrated array is in progress.

In the hybrid IRCCD category most effort has been concentrated on the coupling of photovoltaic detectors. Work is in progress on the development and operation of PbSnTe, InAsSb, InSb and HgCdTe hybrid IRCCDs.

The feasibility of using photoconductive detectors coupled to CCDs through a buffer stage has been demonstrated with HgCdTe. Finally, the pyroelectric hybrid IRCCD is receiving renewed attention for low-cost, low-performance room temperature applications.

In conclusion, it is felt that the IRCCD field has developed at an extremely rapid pace in the past few years from the initial conceptual phase into the feasibility and operation stage. At this point in the development of the field a considerable number of device options are still being pursued for various applications.

REFERENCES

BOYLE, W.S., and G. E. Smith, 1970, "Charge Coupled Devices", Bell System Technical Journal, Vol. 49.

BURKE, H.K., and G. J. Michon, 1976, "Charge-Injection Imaging: Operating Techniques", Proceedings IEEE, Vol. ED-23, p. 189.

ERB, D.M., and K. Nummedal, 1973, "Buried Channel CCD for Infrared Applications", Proceedings of the 1973 CCD Applications Conference, p. 157.

KIM, J.C., 1975, "InSb MIS Technology and CID Devices", Proceedings 1975 CCD Applications Conference, p. 1.

KIM, J.C., et al., 1977, "Development of Indium Antimonide Arrays", Final Report, Contract No. N00173-76-C-0128.

KOHN, E.S., 1976, "A Charge Coupled Infrared Imaging Array with Schottky Barrier Detectors", Proceedings IEEE, Vol. ED-23, p. 207.

NELSON, R.D., 1974, "Accumulation-Mode CCD", Applied Physics Letters, Vol. 35, p. 568.

STECKL, A.J., and T. Koehler, 1973, "Theoretical Analysis of Directly Coupled 8-12 μm Hybrid IRCCD Serial Scanning", Proceedings of the 1973 CCD Applications International Conference, p. 247.

STECKL, A.J., et al., 1975a, "Applications of CCDs to Infrared Detection and Imaging", Proceedings of the IEEE, Vol. 63.

STECKL, A.J., 1975b, "Low Temperature Silicon CCD Operation", Proceedings of the 1975 CCD Applications International Conference, p. 383.

STECKL, A.J., 1975c, "Injection Efficiency in Hybrid IRCCDs", Proceedings 1975 CCD Applications Conference, p. 85.

STECKL, A.J., 1976, "Infrared CCDs", Infrared Physics, Vol. 16, p. 65.

STECKL, A.J., 1977, "Infrared Detector Properties of the p-p PbS-Si Heterojunction", Proceedings IRIS Detector Meeting, Colorado Springs, Colorado.

DISCUSSION

Sirieix

What is your opinion about large integrated infra-red arrays used in starring mode?

Author's Reply

Certainly, that seems to be the long-term goal. To obtain large-scale IR integration in the near-term, we must very carefully choose a device approach that will not suffer from the twin problems of the high background photon flux and response non-uniformity. In the long run, I believe the only solution lies in integrating on chip more sophisticated signal processing along with the read-out function.
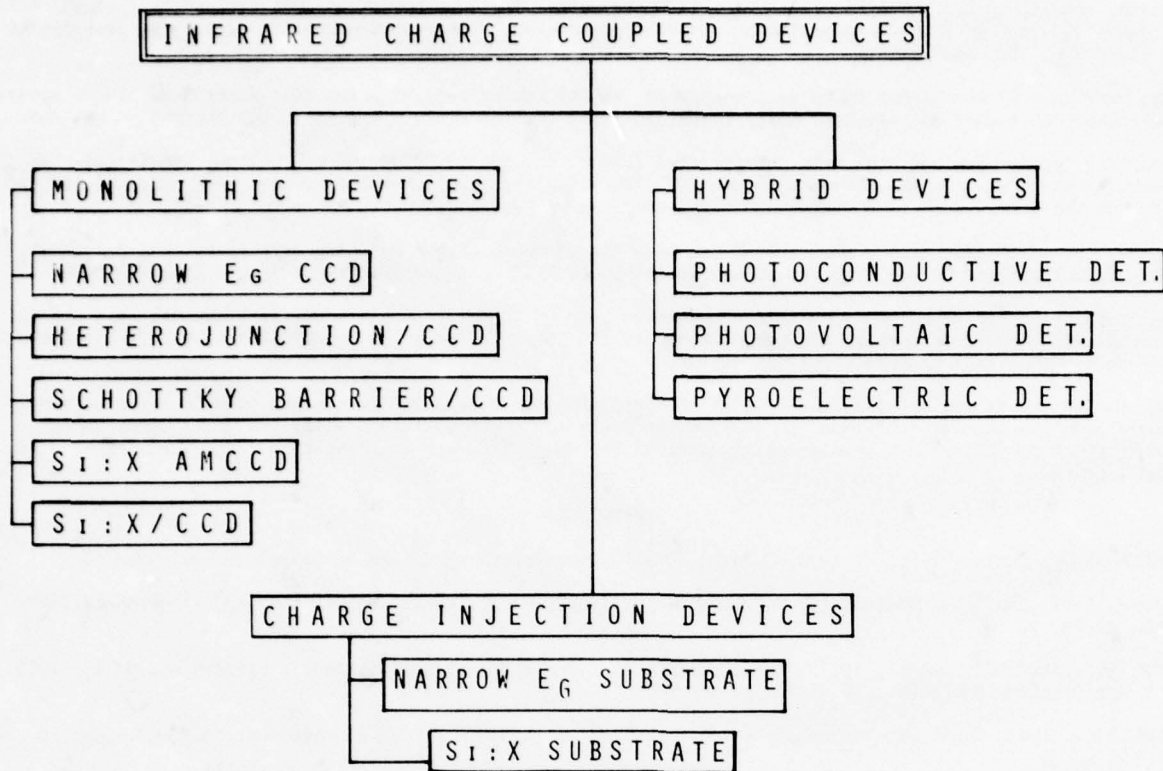
Fig. 1 Classification of Infrared Charge Coupled Devices.
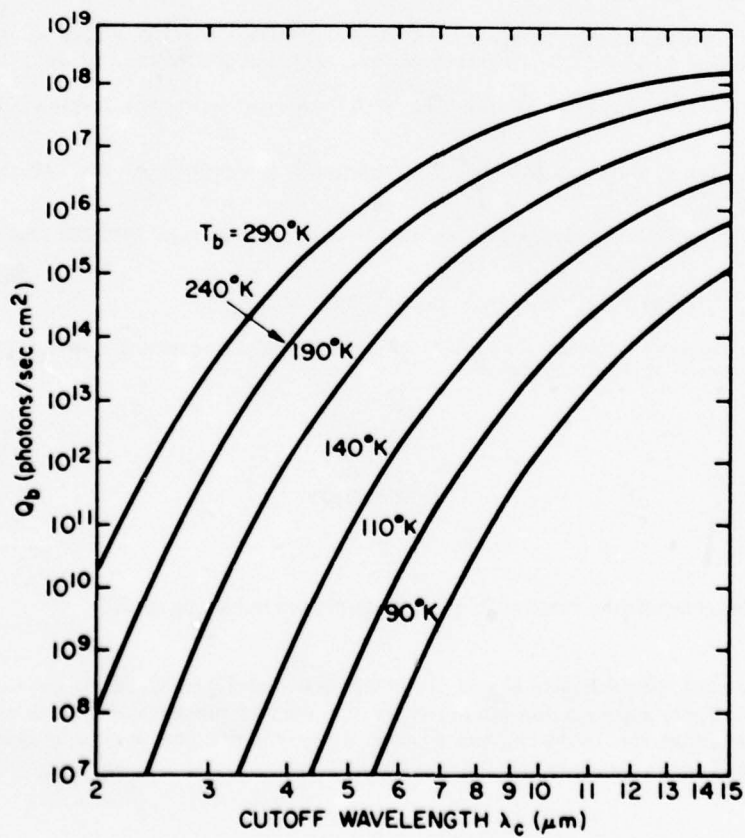
Fig. 2 Photon Flux Density vs. Wavelength.

TEMPERATURE CONTRAST RATIO:

$$\frac{\Delta N_\lambda}{N_\lambda} = \frac{0.16}{\lambda} \frac{\Delta T}{[1 - e^{-0.478/\lambda}]}$$

WHERE λ IN μM
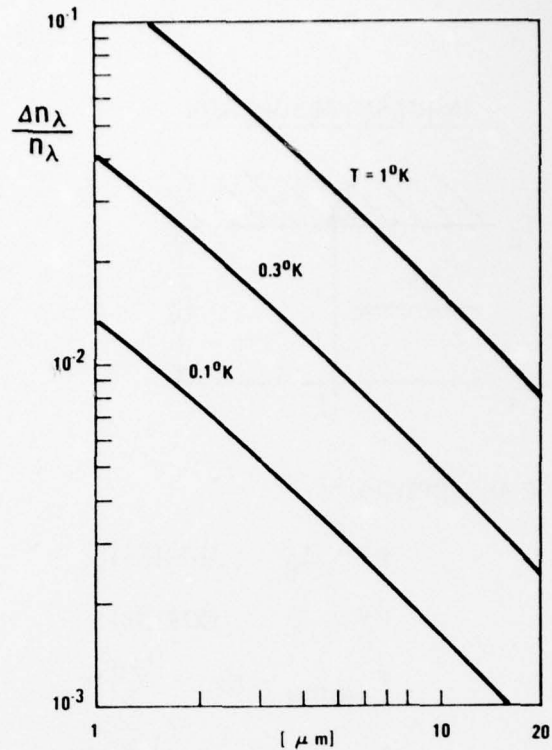
Fig. 3 Contrast vs. Cut-Off Wavelength.
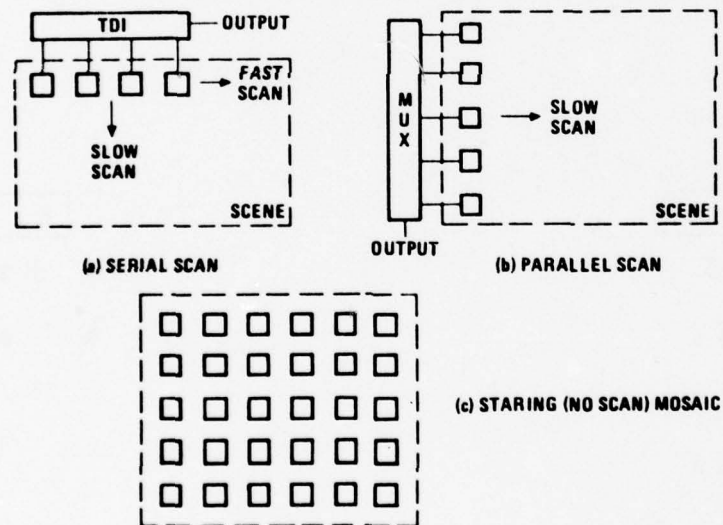
Fig. 4 IRCCD Focal Plane Organizations
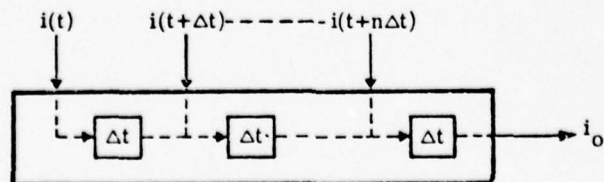
Fig. 5 CCD Time-Delay-and-Integration.

INTRINSIC ABSORPTION

EXTRINSIC ABSORPTION

FOR ABSORPTION

$$h\gamma \geq E_G \qquad \text{INTRINSIC}$$

$$h\gamma \geq E_i \qquad \text{EXTRINSIC}$$

$$E_{PHOTON} = h\gamma = \frac{hc_0}{\lambda}$$

$$E_{PHOTON} (eV) = \frac{1.24}{\lambda (\mu m)}$$
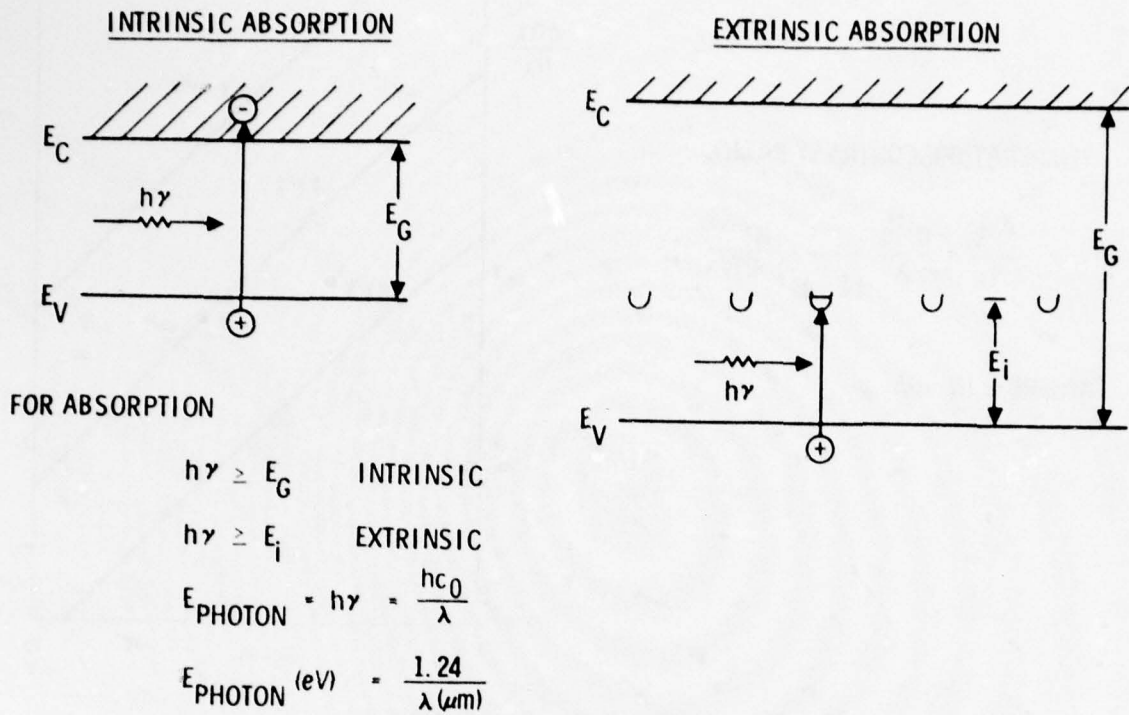
Fig. 6 Photon Absorption Mechanisms.

$\boxed{3 - 5 \, \mu M}$

- $P_B S$
- $P_B S_E$
- $P_B T_E$
- $I_N S_B$
- $I_N A_S S_B$
- $H_G C_D T_E$

$\boxed{8 - 12 \, \mu M}$

- $H_G C_D T_E$
- $P_B S_N T_E$

Fig. 7 Monolithic IRCCD: Narrow Gap Semiconductors.

DEVICE PARAMETERS	MATERIAL PARAMETERS															OPERATION PARAMETERS			
	τ	μ	S_o	x_D	L_D	N_I	N	N_{SS}	ε_S	ε_I	α					T	Q_B	V_G	L
STORAGE TIME	*		*	*	*	*	*	*		*	*					*	*	*	
TRANSFER TIME		*			*	*	*		*	*						*		*	*
QUANTUM EFFICIENCY			*	*		*					*								
DYNAMIC RANGE									*									*	
CROSS - TALK					*					*									*

τ = LIFETIME $\qquad$ L_D = DIFFUSION LENGTH $\qquad$ ε_I = INSULATOR FOR DIELECTRIC CONST.

μ = MOBILITY $\qquad$ N_I = INTRINSIC CARRIER CONC. $\qquad$ α = ABSORPTION COEFFICIENT

S_o = SURFACE RECOMBINATION VELOCITY $\qquad$ N = DOPING DENSITY $\qquad$ T = TEMPERATURE

x_D = DEPLETION WIDTH $\qquad$ ε_S = SUBSTRATE DIELECTRIC CONST. $\qquad$ Q_B = PHOTON FLUX DENSITY

L = GATE LENGTH $\qquad$ N_{SS} = SURFACE STATE DENSITY $\qquad$ V_G = GATE VOLTAGE

Fig. 8 Monolithic IRCCD: Device-Material Interaction.
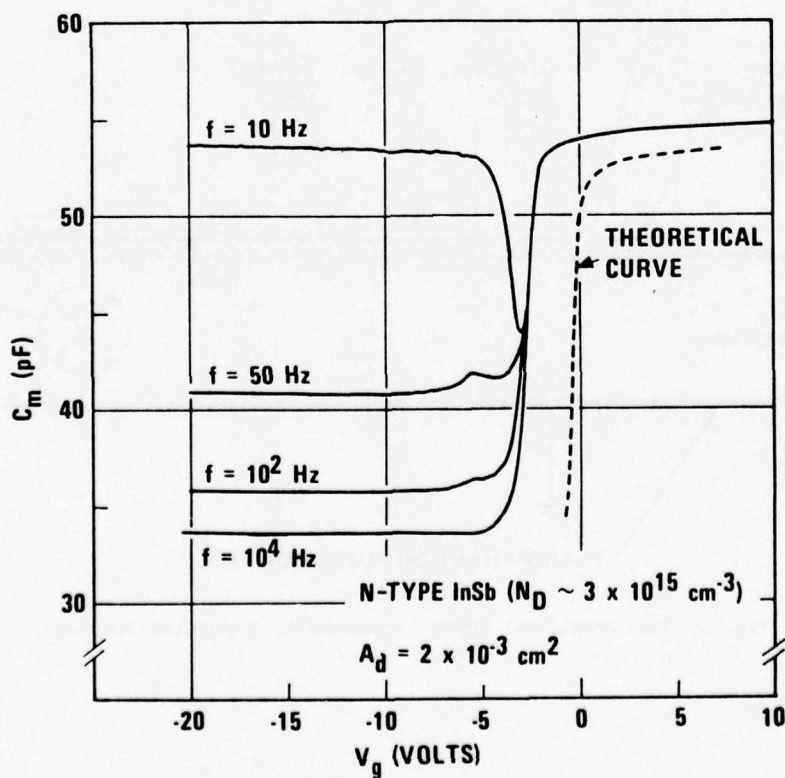
Fig. 9 InSb MIS Structure: C-V Characteristics.

	E_i(eV)	$\lambda_c (\mu M)$	Si: B	E_i(eV)	$\lambda_c (\mu M)$
Si: P	0.045	27.6	Si: B	0.04385	28.2
Si: As	0.0537	23.1	Si: Al	0.0685	18.2
Si: Sb	0.043	28.8	Si: Ga	0.0723	17.2
Si: Bi	0.0706	17.6	Si: In	0.1554	8.0
Ge: Cu	0.04	31.0	Ge: Zn	0.03	41
Ge: Ag	0.13	9.5	Ge: Cd	0.05	25
Ge: Au	0.16	7.7	Ge: Hg	0.09	13.8

Fig. 10   Impurity Ionization Energies and Corresponding Cut-Off Wavelengths of Extrinsic Si and Ge Photoconductors.
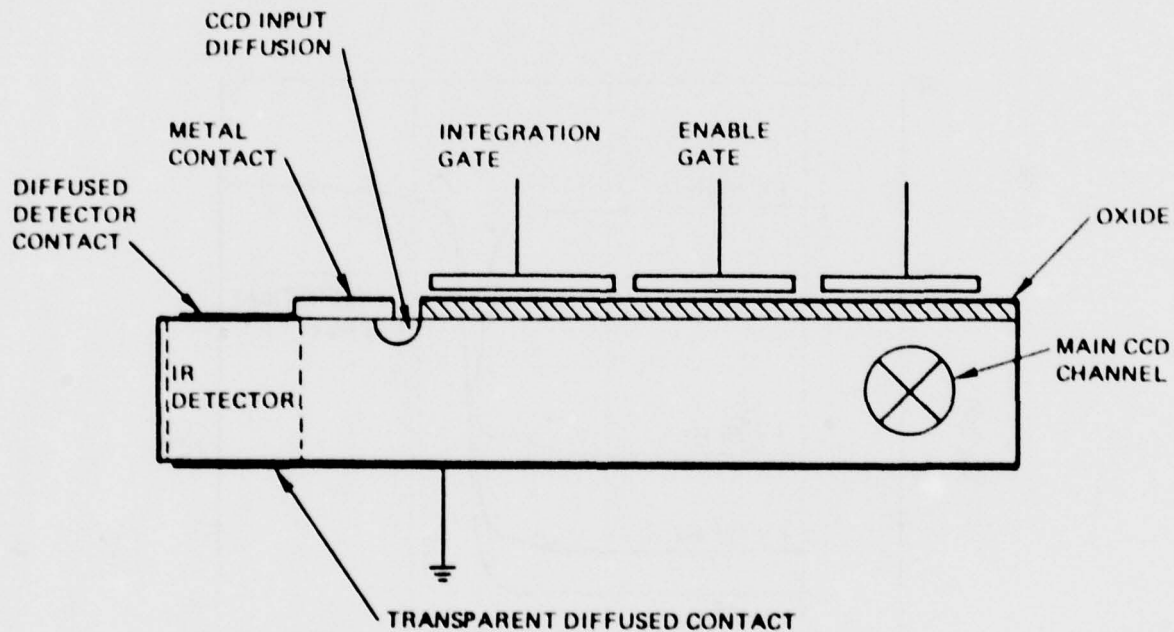
Fig. 11   Si:X Monolithic IRCCD: Horizontally Integrated Structure.
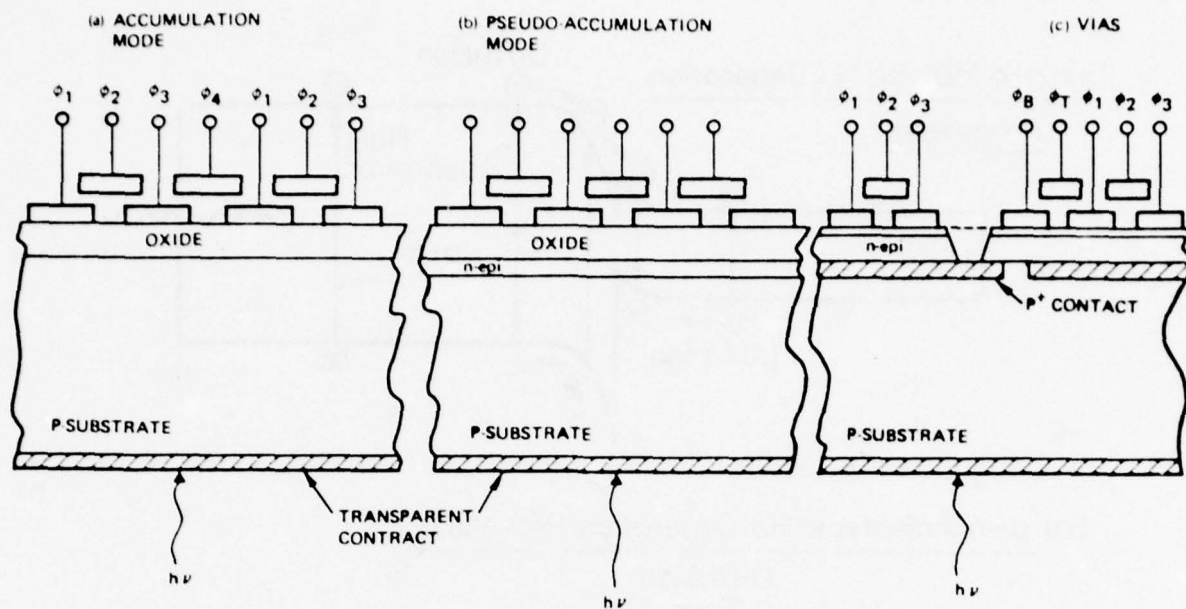
Fig. 12 Si:X Monolithic IRCCD: Vertically Integrated Structures.
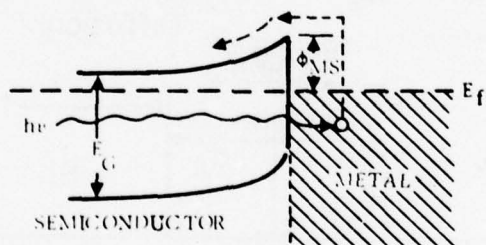
Fig. 13 Band Diagram of Schottky Barrier Detector.

Fig. 14 Photogeneration Processes in the PbS-Si Heterojunction Detector.

IMAGE
SCAN →

IR DETECTOR CHIP

Si CCD CHIP

CHARGE
TRANSFER →
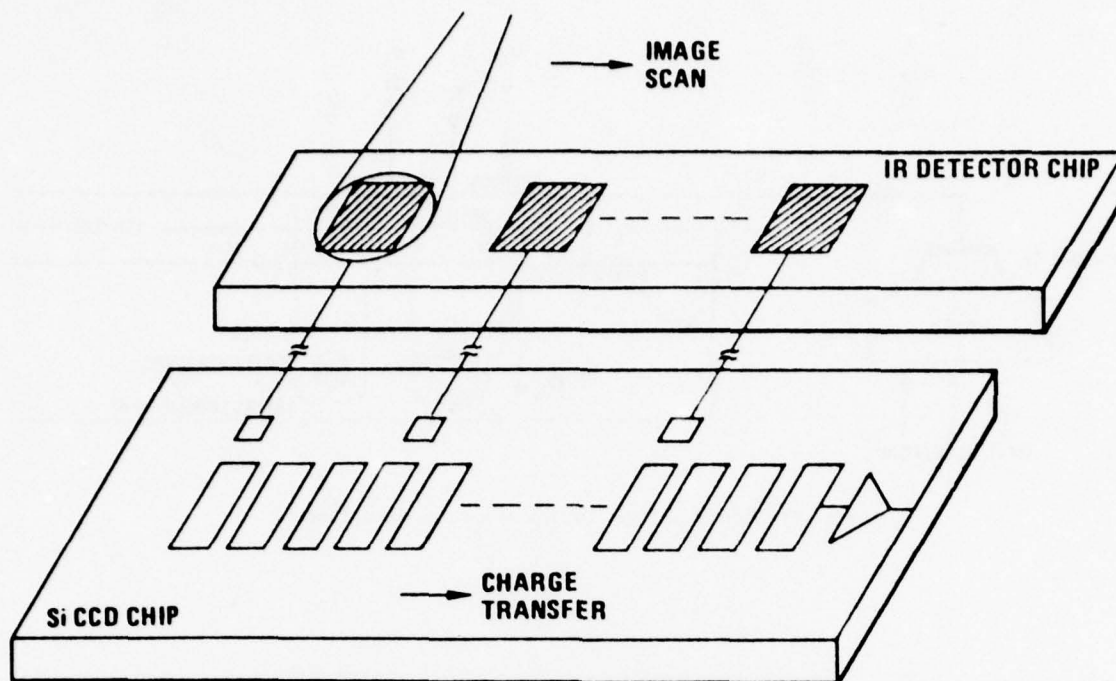
Fig. 15 Hybrid IRCCD: TDI Serially scanned Operation.

IR DETECTORS

DIRECT INJECTION
INPUT

GATE INPUT
MODULATION

V_{ID}

BUFFER STAGE
INPUT

V_{IG}

CCD SHIFT REGISTER
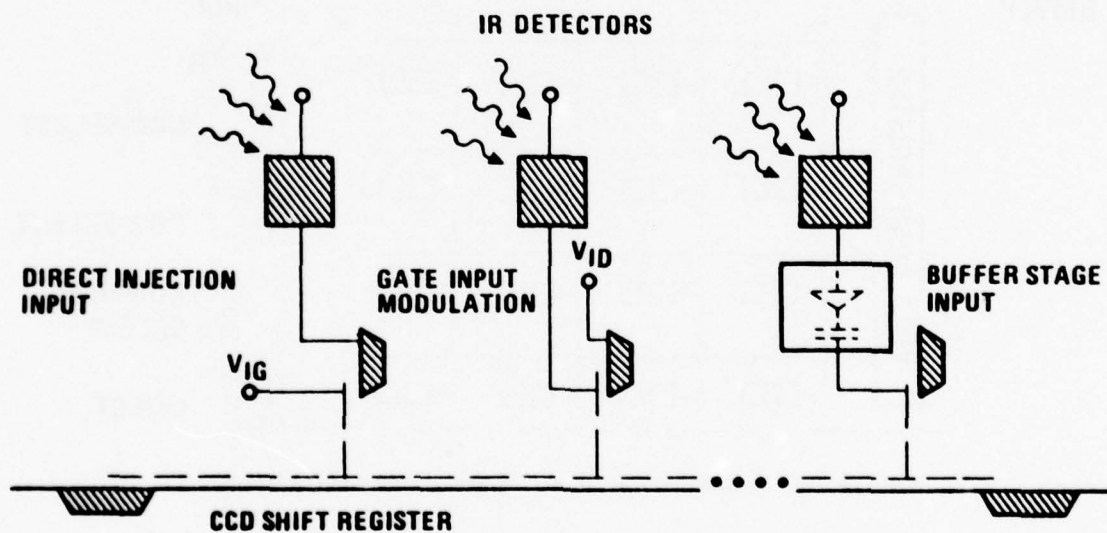
Fig. 16 Hybrid IRCCD: Input Coupling Schemes.
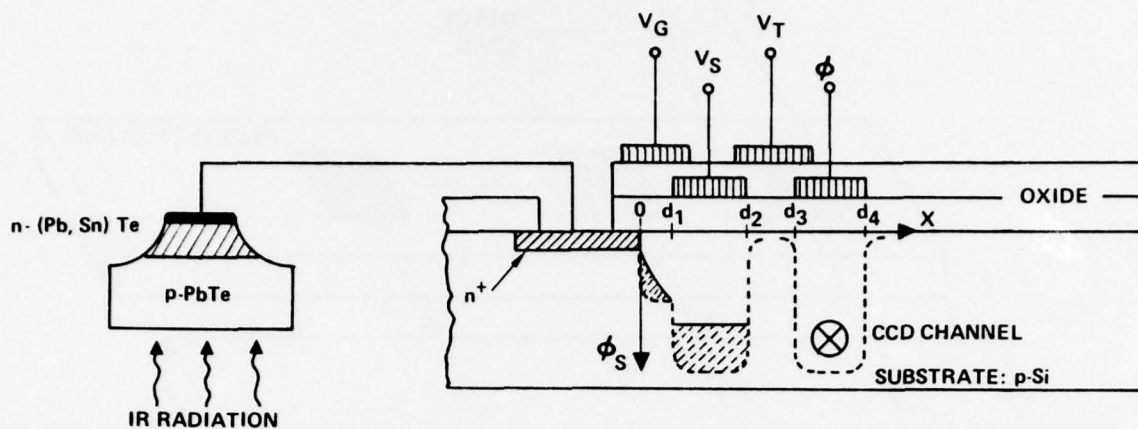
Fig. 17   Hybrid IRCCD: Direct Injection Interface.
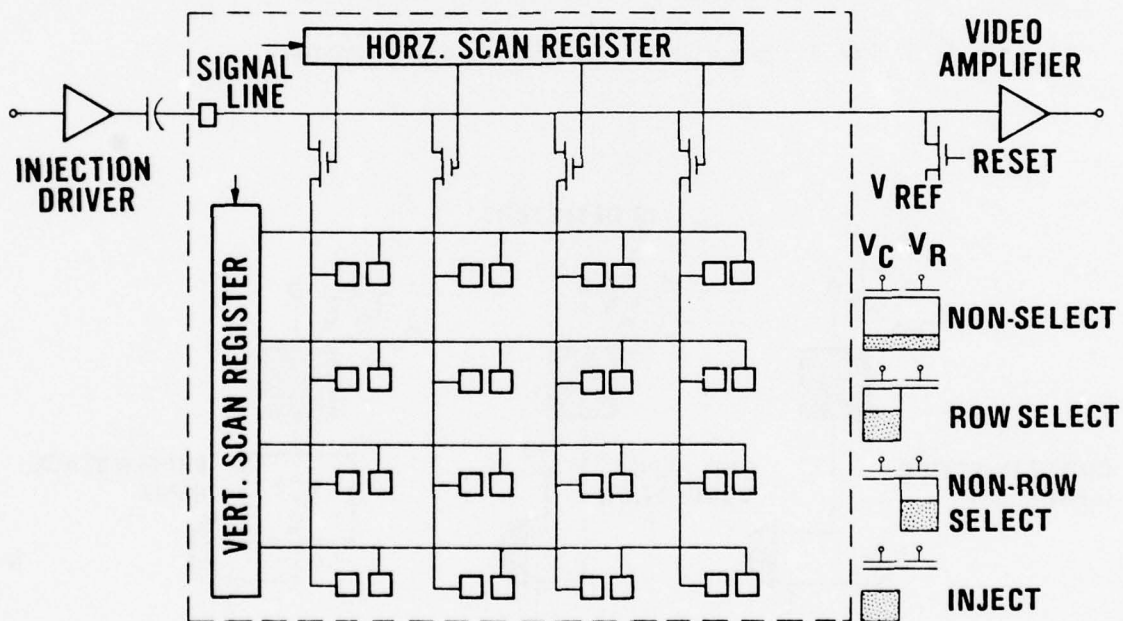
Fig. 18   CID Two-Dimensional Array: Organization and Operation.

A. NARROW GAP SEMICONDUCTOR DEVICES

    1. $InSb$ → MIS, MOSFET, CCD DEMONSTRATED
    2. PbS
    3. $PbTe$
    4. $HgCdTe$ → MIS WORK IN PROGRESS

B. EXTRINSIC SILICON DEVICES

    1. $Si:Ga$
    2. $Si:Al$ → IRCCD DEMONSTRATED. WORK IN
    3. $Si:In$ PROGRESS ON VARIOUS SIZE ARRAYS

C. SCHOTTKY BARRIER DEVICES

    $Pd-Si$ → IRCCD DEMONSTRATED

D. HETEROJUNCTION DEVICES

    $PbS-Si$ → DETECTOR DEMONSTRATED, IRCCD WORK IN PROGRESS

Fig. 19 Monolithic IRCCD Work in Progress.

1. PHOTOVOLTAIC DETECTORS

    * $PbSnTe$
    * $HgCdTe$
    * $InSb$ FEASIBILITY PROVEN. WORK
    * $InAsSb$ IN PROGRESS ON VARIOUS SIZE ARRAYS

2. PHOTOCONDUCTIVE DETECTORS

    * $HgCdTe$ - FEASIBILITY PROVEN

3. PYROELECTRIC DETECTORS

    * $LiTaO_3$
    * S B N   FEASIBILITY WORK IN PROGRESS
    * T G S

Fig. 20 Hybrid IRCCD Work in Progress.

CCPD--THE OPTIMUM SOLID-STATE LINE SCANNER

Hsin-Fu Tseng and Gene P. Weckler
Reticon Corporation
910 Benicia Avenue
Sunnyvale, California 94086

ABSTRACT

A new family of optimized solid-state line scanners has been developed and is now available to the system designer. These devices offer many advantages for applications requiring high sensitivity, low noise, and spectral purity. The devices to be described employ photodiodes as the detector elements, thus assuring maximum quantum efficiency and a broad, smooth spectral response, both characteristic of a good silicon photodiode. A charge-transfer device is employed to provide the low-noise readout. This combination of photodiodes and analog shift register results in the optimum solid-state line scanner. In addition, these devices contain anti-blooming circuitry as well as output buffer amplifiers.

The charge-coupled photodiode array is intended specifically for applications requiring spectral purity, i. e., real time spectrophotometers and multispectral scanners. The smooth spectral response further provides improved uniformity to monochromatic illumination. These, as well as other advantages of this architecture, are discussed and examples of performance are presented.

1. INTRODUCTION

Evolution has produced several solid-state image sensors, each possessing a different architecture. Most of these architectures can be broken down into combinations of four basic building blocks. This paper will present a review of the most common architectures and discuss in detail the new architecture which results in the optimal solid-state line scanner.

The solid-state line scanner takes advantage of the highly developed silicon integrated-circuit technology. The mechanism of detection is based on the absorption by silicon of photons within an energy range of 1.1 ev to about 6 ev. This corresponds to a wavelength range of 0.2μ to 1.1μ. When a photon is absorbed, it generates an electron-hole pair. If we are to detect this electron-hole pair, the components must be separated. This separation is normally accomplished by the depletion region of a p-n junction or the depletion region induced by applying the appropriate voltage to an MOS capacitor. This latter depletion region is also referred to as a potential well. In either case, the electron and hole are separated and the charge equivalent to one electron will then appear on the depletion-region capacitance. Let us briefly compare these two basic detection mechanisms. The internal quantum efficiencies can for all practical purposes be assumed to be the same for both mechanisms, i. e., the efficiency of collecting photo-generated electron-hole pairs. The main difference is in the external quantum efficiency of the two mechanisms. Figure 1 shows the basic structure of these two detectors. The external quantum efficiency of the diffused photodiode suffers minimum losses due to the presence of only two interfaces between materials of different refractive indices, i. e., Air-SiO_2 interface and SiO_2-Si interface. The thickness of the SiO_2 is such that the modulation of the spectral response of the diffused photodiode is negligible. It is apparent from Figure 1 that for the field-induced detector an additional two interfaces are present to introduce losses. Furthermore, the transparent electrode is not really transparent since it is usually polysilicon. Since silicon is absorptive, some of the incident photons are absorbed in this layer. This is particularly true for the short wavelength or the blue end of the spectrum. The use of exotic metallic materials has resulted in field plates that are more transparent over the spectral range of interest than is polysilicon; however, these materials are foreign to standard integrated-circuit technology.

The absorption of incident photons in this quasi-transparent layer can be reduced by making the layer thin. This, however, usually results in modulation of the spectral response due to interference patterns set up in the electrode layer and in the silicon dioxide layer between the electrode layer and the silicon since it too must be thin to insure reasonable operating voltages. Furthermore, the thickness of these films are subject to normal processing variations. It is, therefore, difficult to insure reproducability of sensitivity, uniformity, or spectral response. It is apparent that the diffused photodiode is a far superior detector, possessing the following advantages: (1) external quantum efficiency approximately three times that of the quasi-transparent electrode employing polysilicon; (2) full spectral response extending from 0.2u to 1.1μ; and (3) a relatively smooth spectral response not subject to process variations.

Having now detected internally the absorption of a photon, it is necessary that this information be made available at a terminal. Here lies another of the principal differences in the design of solid-state line scanners. Figure 2 shows schematically two approaches used to interrogate and read out the individual picture elements of a line scanner. Each approach uses a shift register to read out the information stored on each individual photo-sensitive element or pixel. In the first case a digital shift register is used to sequentially access a transfer switch which connects individual pixels in turn to a common terminal. This approach has the definite advantage that digital shift registers and multiplex switches have been highly developed, use standard MOS processes, and are relatively easy to implement. The performance of this readout technique is dependent on both the total number of multiplex switches and on the uniformity of the multiplexing function, i. e., ideally each multiplex switch and its drive should be identical. Non-uniform

multiplexing results in a fixed-pattern modulation which is superimposed on the video information from the pixels. Differential signal-processing techniques have recently been incorporated which have reduced the fixed-pattern component to a negligible level. The fixed pattern has been reduced to the point where the total number of multiplex switches is now of practical significance. The random noise depends directly on the size of the output capacitance which in turn is a function of the number of multiplex switches connected to the output line. In the majority of applications, particularly those for which the solid-state line scanner serves as an input to a machine, random noise does not appear to be a practical problem. The level of the random noise, however, does set a basic limit to the minimum detectable illumination level that can be detected, and in doing so, must be considered in any design.

The second approach, shown in Figure 2, also employs a transfer switch (really an adjustable barrier) for each pixel; however, all pixels are sampled simultaneously, thus transferring all the information in parallel into an analog shift register. This information is then clocked to an output terminal at the end of the analog shift register. The analog shift register has been highly developed over the past few years. Charge-transfer devices, both bucket brigade and charge-coupled, can now be made with transfer efficiencies exceeding 0.9999 at megahertz clocking rates; therefore, the initial problems of shading and loss of resolution are no longer a serious problem.

Figure 3 shows four architectures that may be implemented using the building blocks described above. Let us begin by examining each structure. The first structure to be discussed uses photodiodes as the detectors and a digital shift register to sequentially interrogate these diodes and is depicted in the figure as Combination A. This structure operates in the charge-storage mode and is commonly referred to as a self-scanned photodiode array. To obtain line storage requires a single multiplex switch connected to each photodiode, thus making possible high-density linear arrays which possess all the advantages of the photodiode detector. For linear arrays, this architecture is perfectly adequate for realizing long high-density arrays. Linear arrays 2000 pixels in length with pixels as close as 15 micron centers have been available for some time. This architecture, however, has reportedly two serious shortcomings. The one most often referred to is the output capacitance, which increases directly with the number of pixels. Its effect is to increase directly the thermodynamic or random noise of the system, thus limiting the minimum number of photons/pixels that can be detected. For arrays of less than about 10^4 pixels the resulting increase in noise due to increasing output capacitance is usually of no practical importance since other sources of noise usually dominate.

It is usually the fixed-pattern noise that has limited the number of grey levels that are discernable. The fixed pattern noise has two origins. Slight differences in the shift register and multiplex switches have been in the past a source of a fixed-pattern noise discernable primarily at lower levels of illumination. This, however, has been virtually eliminated by differential techniques and is, therefore, much less of a problem now as compared to earlier arrays using this architecture. Another source of fixed-pattern noise results from the pixel-to-pixel variation in photosensitivity. This sensitivity variation results in an uncertainty that also limits the number of grey levels that can be discerned, but is a fundamental problem resulting from the physics and not from the architecture.

The second architecture to be discussed is commonly referred to as a charge-coupled device which uses the field-induced photo-detector as the pixel and the analog shift register to shift the information from the pixel to the output terminal and is depicted by Combination B. Depending on the particular criteria employed, the performance of these structures has ranged from adequate to excellent. As a result of the very low output capacitance and the elimination of sequential sampling with multiplex switches, both the thermodynamic and the fixed-pattern noise in the dark are exceptionally low. This, however, is offset by the resulting non-uniformity that prevails under illumination. This non-uniformity is a result of the variations in film thicknesses that occur in fabricating the field-induced photo-detector added as well as those non-uniformities that are always present in the bulk silicon. Since the reflectivity as well as the absorption depends on the relative thicknesses of several films, a compromise must be made between spectral response, quantum efficiency, and the non-uniformity. Normal process variations make reproduction of consistent parameters over a period of time somewhat more difficult than for the simpler diffused diode structure.

The third structure to be discussed, shown as C in the figure, combines the field-induced photo-detector with the digital shift register in an effort to obtain higher density with an existing technology. As initially conceived, this structure exhibited excessive uncontrolled blooming, less sensitivity than the photodiode, spectral variations, excessive non-uniformity, fixed patterns in the dark resulting from digital sampling, and an extremely large output capacitance. Most of these difficulties are now under control; however, the technology is no longer standard, requiring an exotic metal/silicon-gate MOS process on an epitaxial substrate. Furthermore, a double-sampling technique must be used to process out the fixed-pattern noise resulting from the sequential sampling of multiplex switches and the KTC thermodynamic noise associated with resetting the output capacitance. As a result of employing this more complicated signal processing technique, the inherent forms of signal contamination are eliminated, and good low-level performance is obtained.

The final structure to be assembled from the set of building blocks is shown in the figure as D. This structure uses photodiodes with all their inherent advantages, i. e., spectral purity and high-external quantum efficiency, combined with an analog shift register for readout. This combination possesses all the advantages of the photodiode detector with those of the analog shift-register readout. The discussion to follow will describe the practical realization of a linear array employing this architecture. For lack of a more descriptive acronym, let us refer to this architecture as the charge-coupled photodiode array (CCPD array).

2. REALIZATION OF THE CHARGE-COUPLED PHOTODIODE ARRAY

The architecture of the charge-coupled photodiode array has been realized using standard production MOS processes.

Figure 4 shows schematically the basic structure of the array. The sensitive region consists of a row of photodiodes. All photodiodes simultaneously dump their charges into an analog shift register which, in turn, transfers the information to the output terminal. Actually, two analog shift registers are used with alternate photodiodes being dumped into the same readout register. This not only doubles the data rate, but also results in a "full wave" or zero-order sample-and-hold output.

3. FUNCTIONAL DESCRIPTION

As stated earlier, the CCPD is a linear diode array with two readout shift registers. The shift registers are n-channel four-phase (4ϕ) surface-channel CCD fabricated using a double poly-silicon n-channel technology. The center-to-center distance of the photodiodes is 16μm, and the channel width of the CCD shift register is approximately 160μm. A simplified schematic is shown in Figure 4. The sensing region consists of a row of diffused photojunctions, connected interdigitally each to a separate diffused island, as in Figure 5. On one side of the diffused island is an anti-blooming gate V_R which controls the excess signal charge and drains it into a sink voltage V_{DD}. On the other side of the diffusion island is a transfer gate ϕ_t which interfaces the signal charge into the CCD register. Each CCD shift register is provided with a low-noise fill-and-spill input and a charge integrator at the output node. The output signals of both CCD registers are multiplexed off-chip to obtain the combined video signal. The whole chip is covered by a second metal layer except for a 16μm aperture above the sensing diodes.

For normal operation, the device requires a substrate bias of -5V, and the CCD register is driven by a set of standard 4ϕ clocks swinging from 0V to +15V. The blooming control gate V_R is grounded and the transfer gate clock ϕ_t swings from -5V to +5V. Figure 5 shows the cell layout, and Figure 6 the potential diagram showing the anti-blooming effect of V_R as well as the charge transfer path into the CCD shift register. When the transfer gate ϕ_t is turned on, signal charge is transferred into the shift register and the sensor diode is reset to a level of "$V\phi_t - V_T$" where V_T is the threshold of the ϕ_t gate. The integration time is the period between two subsequent ϕ_t or transfer pulses as depicted in Figure 7. After the sensor diode is reset, the signal charge starts to build up as the photons are absorbed at each sensor site. When the signal charge reaches the level set by the blooming control gate V_R, the excess charge will spill over the V_R barrier and sink into V_{DD}. The barrier between each sensor diode is at the substrate voltage -5V. There are 5V margins between the substrate barrier and the blooming control barrier. Therefore, the anti-blooming effect is excellent. By proper clocking, the blooming control gate also can be used as an exposure[1] or contrast[2] control.

While the sensing diode is integrating the incident photon flux, the shift register delivers the previous signal charge to an output node, where a charge integrator converts it pixel by pixel to an output voltage. Since all the signal charge is differentiated from a reference level set by the transfer gate ϕ_t, no fixed pattern noise is to be expected. The input port at the input end continuously supplies a background charge to minimize the trapping effect, thus optimizing the transfer efficiency. This input port also can be used to monitor the transfer efficiency of the shift register by feeding in a square-wave signal and measuring the percentage charge loss of the leading pixel. A wide channel is used to maximize the transfer efficiency. The measured transfer efficiency is better than 0.9999 at 2.5MHz clock rate, which corresponds to a 5MHz data rate at the combined video output. To obtain optimum performance, clock waveshape control of the major clocks, ϕ_1 and ϕ_3, is desirable. Specifically, it is desirable to control overlap and rise and fall times, as illustrated in Figure 8.

4. PERIPHERAL INTERFACE CIRCUIT

Although the device's peripheral circuit employs standard CCD techniques, caution should be used in developing the clocks and output interface circuitry to obtain optimum performance from the device. Therefore, a brief description is given of the interface peripheral circuit which has been employed to measure the device performance. Also, some of the interrelationships between the device and the peripheral circuit are given.

A simplified block diagram is shown in Figure 9. The circuit is separated into three sections: (a) the sensor, (b) the logic control, and (c) the analog signal-processing circuit. The logic control section generates the four-phase clocks and the transfer pulse for the sensor array, along with the sampling and dc restoration pulses for the signal-processing circuits.

The signal processing circuits are composed of the preamplifier, the sampling switch, and the sample-and-hold capacitor, the switch-edge-cancellation (SEC) capacitor, a simple filter, and the buffer amplifier for the output. The odd and even pixels from the array are summed at the preamplifier via the gain-balancing potentiometer. The combined samples are then resampled and held on the holding capacitor for the duration of a single pixel, then processed with dc restoration, filtering, blanking and buffering.

In the resampling process, sampling switch energy induces extraneous charge on the storage capacitor. The switch-edge-cancellation (SEC) trimmer capacitor is adjusted to minimize or cancel the extraneous charge by coupling charge from a negative-going version of the sampling pulse.

The dc restoration circuit restores a ground reference to the sampled-and-held video during the retrace or blanking time of the array.

The simple RC low-pass filter is used to minimize residual switching spikes and to maintain minimum noise bandwidth consistent with the array switching speed.

The array requires four-phase clocks, a transfer pulse, and complementary reset clocks. The clock waveforms have crossing-point and edge-control requirements as illustrated in Figure 8, and the transfer pulse must occur at the proper time while the four-phase clocks are held stationary, as shown at Figure 6. The logic section provides the required sequencing and shaping, with all timing controlled by a basic trigger oscillator operating at twice the desired sample rate or four times the four-phase clock rate. The sample rate is variable from less than 100KHz to more than 4MHz. Ring-counter-connected flip-flops generate two pairs of complementary square waves, staggered by one trigger period, which provide the basic four-phase clock timing. Direct and delayed versions of the major clocks are combined in NOR gates to give the desired high-level crossing.

Three separate but interconnected counters are used to sequence the transfer process, the readout process, and the desired integration period. During the transfer period, the four-phase clocks are stopped with a particular phase relationship. The readout period follows next, and then the video output is blanked for the balance of the integration period and the transfer period.

One of the two readout shift registers has an extra half stage (Figure 4) with the result that even though pixel information is transferred to both registers simultaneously, output occurs alternately, automatically in the correct multiplex sequence. A second effect of this arrangement is that a "blind" pixel time slot appears at the beginning of the readout sequence, with a level suitable for ground reference (dc restoration) for the video processing circuit. The logic control selects this time slot to provide a clamp pulse to the ground-reference circuit. Finally, the logic control provides the sampling pulses for the output sample-and-hold processor.

The analog video portion of the peripheral circuit provides gain balance between the two readout paths, combines the signals, samples and holds the values pixel by pixel, restores a ground reference, provides a simple low-pass filter, and delivers a buffered low-impedance output. As part of the sampler, a transient glitch-balancing circuit largely cancels the capacitive coupling inadvertently present in the sampler switch. This switch-edge-canceller (SEC) introduces equal and opposite capacitive transients to those from the switch.

5.   PERFORMANCE

The performance of the device has been measured with a circuit similar to that described. Typical performance is shown in Figures 10 to 15.

Figure 10 shows the typical silicon photodiode response which is the same response exhibited by the CCPD because each photosite is a silicon diode. This photodiode response differs from that of the field-plate CCD sensor which exhibits interference patterns that cause bandpass ripples in wavelength response; in addition, the latter response rolls off at the band edges with substantial loss of blue wavelength response.

Figure 11 is a typical optical-to-electrical transfer characteristic. The light exposure in μ Joules/cm^2 is plotted on the x-axis and the output in volts is plotted on the y-axis. These measurements have been taken with a 2870°K tungsten source and a reference detector with a flat response over the band 370 to 1040nm. Note: The drive circuits that were used for these measurements are as described in the foregoing circuit description, but the signal was measured directly across a 1.5K resistor as shown in the schematic insert in Figure 11.

Figure 12 is a family of curves showing the reciprocity between light intensity and integration time. These curves have been measured by holding light power constant and varying the integration time. The output is the voltage output normalized to the saturated light output.

It is obvious that longer integration times may be used to detect lower light levels. However, this approach is ultimately limited by leakage current which is integrated along with the photocurrent. At room temperature, dark current will contribute less than 1% of a saturated signal for integration times up to about 30msec (corresponding to approximately 60KHz scan rate for a 1728-element array). Figure 13 shows dark scans at 1MHz and at 100KHz. The fact that dark signal remains negligible even at the lower scan rate is characteristic of the low-dark-current capability inherent in the CCPD technology.

Figure 14 is an oscilloscope photograph illustrating the full-wave boxcar-type video output with a single cell illuminated by a 1/2 mil diameter light spot. This measurement was taken with the complete signal processing circuit previously described.

Figure 15 illustrates performance of the on-chip video output buffers. The buffer function is to convert the collected video charge into a corresponding output voltage from a relatively low-impedance generator.

6.   CONCLUSION

The basic building blocks of the optimum solid-state image sensor have been available for some time, but until now they have not been optimally combined. Not only does this structure excel in its electro-optical performance, it also provides electrical versatility as demonstrated by its variety of operating modes.

REFERENCES

1.   Wen, D. D. et. al., May, 1975, "The Latest in CCD Image Sensor Technology", presented at Semicon/West.

2.   Kohn, E. S. et. al., 1975, "Infrared Imaging with Monolithic CCD-Addressed Schottky-Barrier Detector Array", proceedings of the 1975 International Conference on the Application of Charge-Coupled Devices

COMPARISON OF PHOTODEFECTOR STRUCTURES
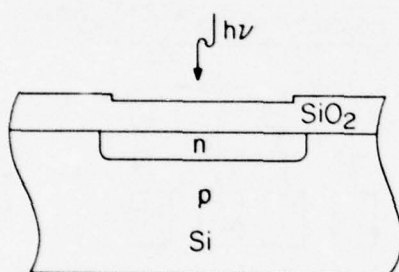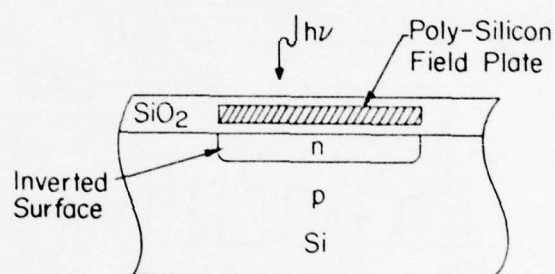
DIFFUSED PHOTODETECTOR   FIELD INDUCED PHOTODETECTOR

Figure 1.   Two Basic Photodetector Structures
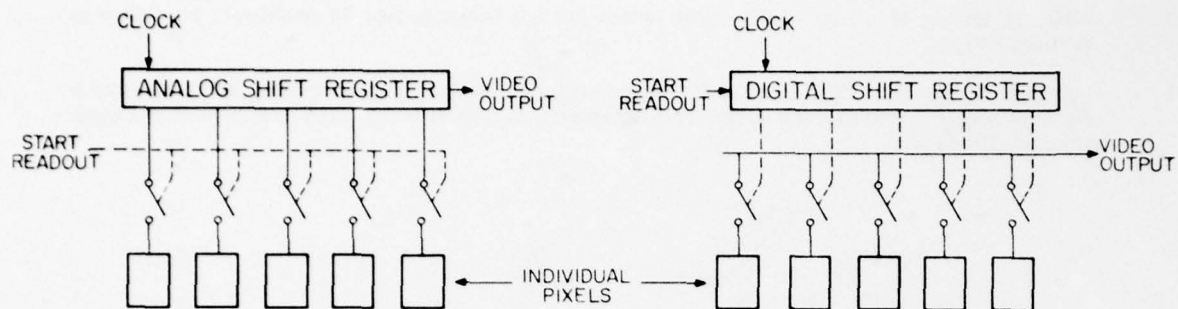
Figure 2.   Techniques for Interrogating and Reading-out Picture Elements

DIFFUSED PHOTODETECTOR   FIELD INDUCED PHOTODETECTOR
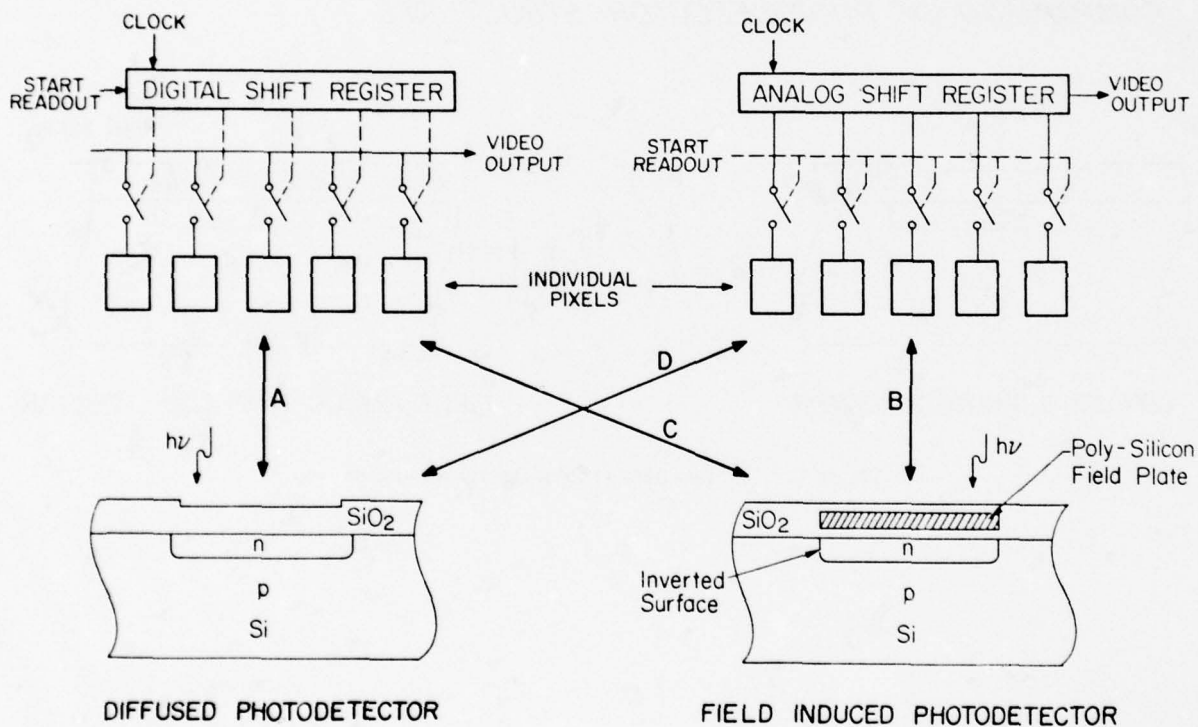
Figure 3.   Four Basic Architectures of a Solid-state Image Sensor
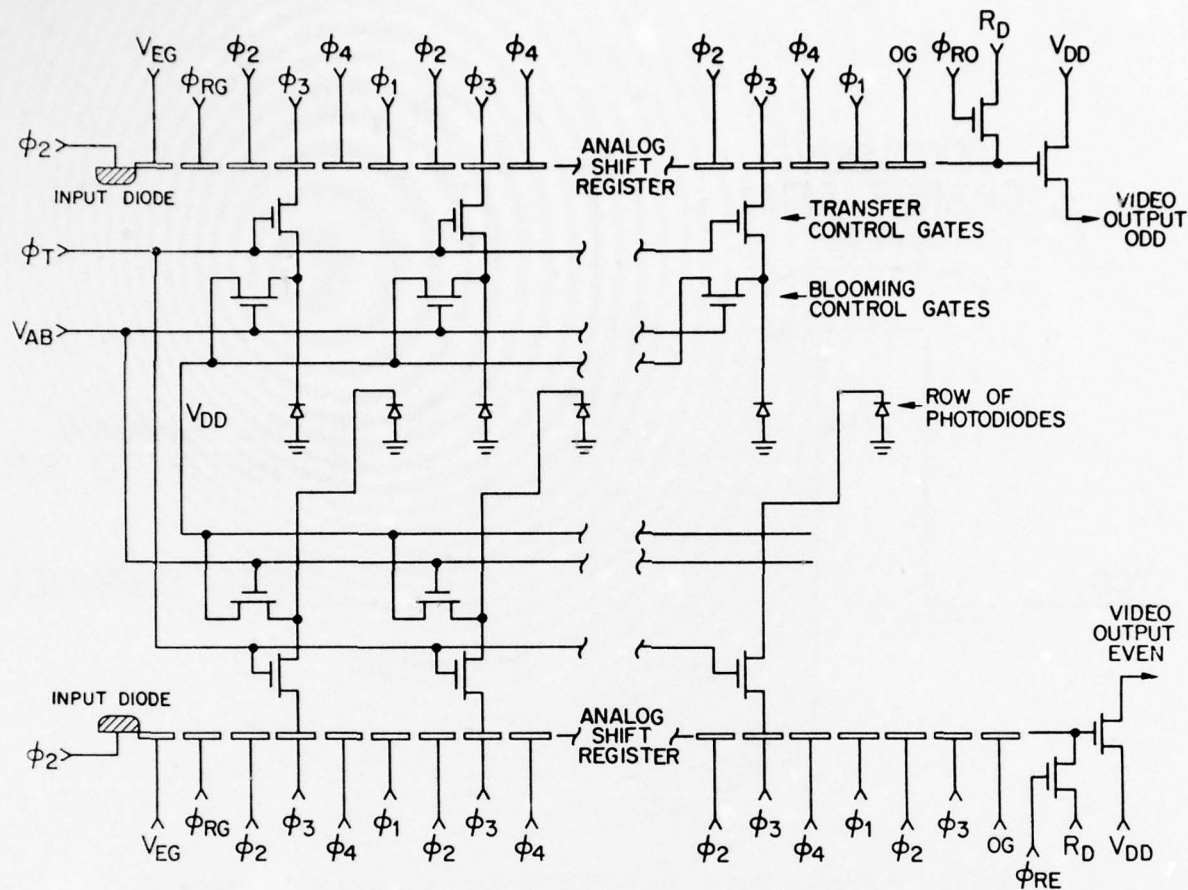
Figure 4.   Simplified Schematic Diagram of the CCPD Linear Image Sensor
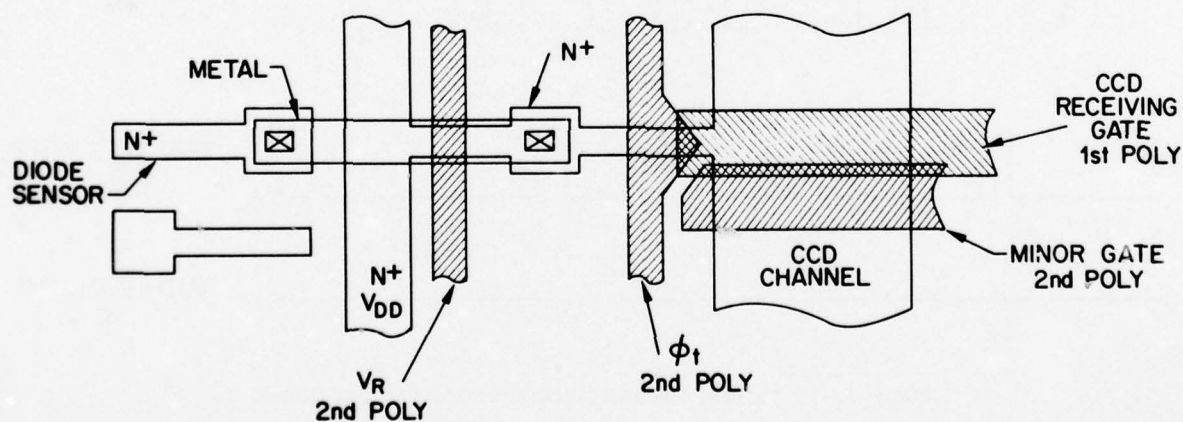
Figure 5.   Details of Construction Showing the Transfer and Anti-blooming Controls
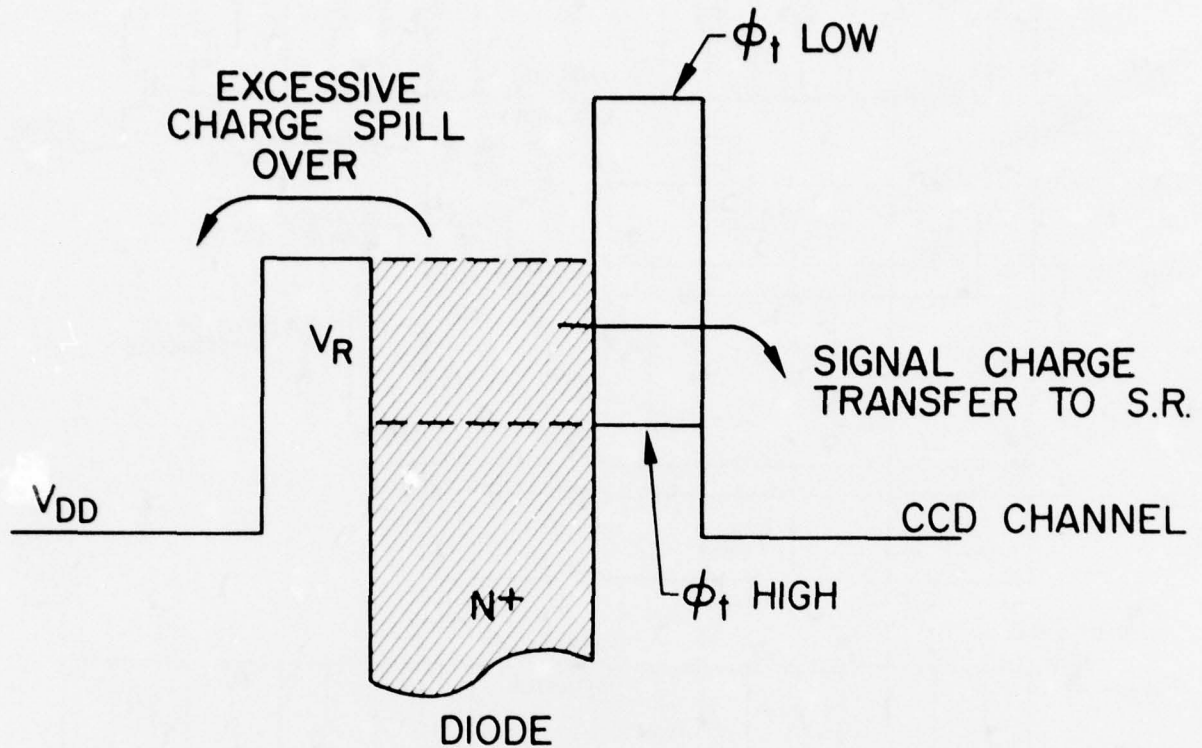
Figure 6.   Potential Diagram Illustrating Transfer and Anti-blooming Properties
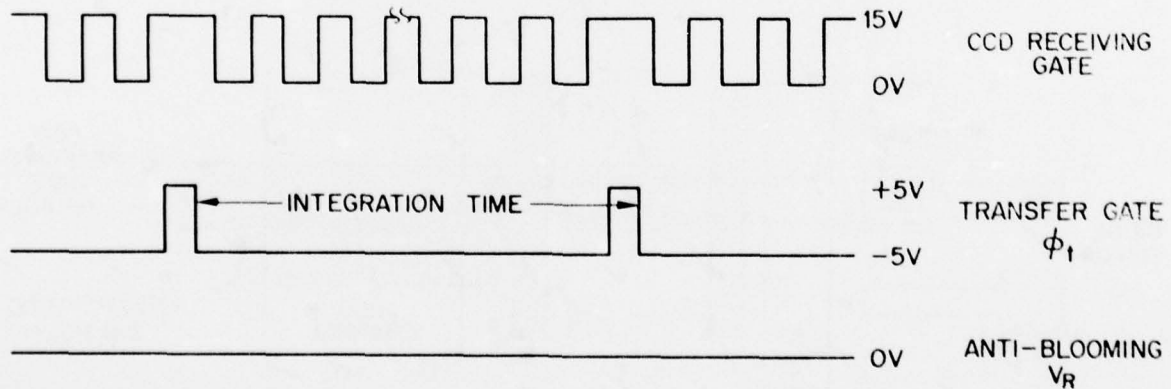
Figure 7.   Timing Diagram Illustrating Transfer and Output

t_r = RISE TIME $\}$ $t_r < 15$ nsec
t_o = OVER LAP TIME $0 < t_o < 10$ nsec
t_f = FALL TIME $\}$ $20 < t_f < 40$ nsec

(a)   (b)

Figure 8.   a) Illustration of Four-phase Clock Timing

                 b) Expanded Scale Showing Details of ϕ_1 and ϕ_3

Figure 9.   Block Diagram for the Peripheral Control Circuit
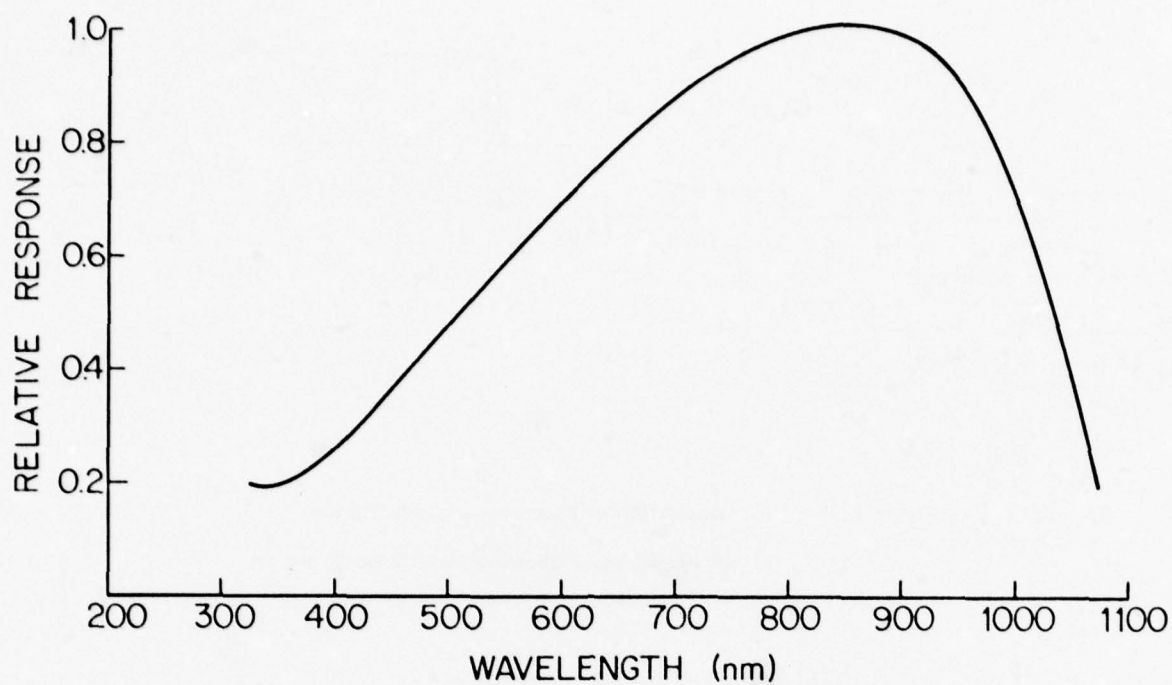
Figure 10.   Typical Spectral Response of a Diffused Diode
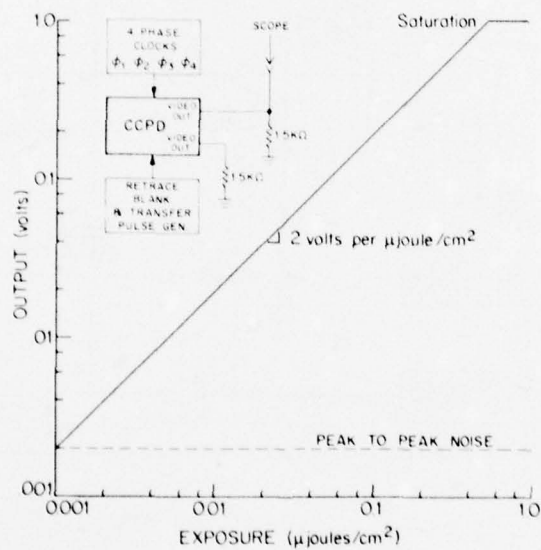
Figure 11.   Optical-to-electrical Transfer Characteristic for the CCPD
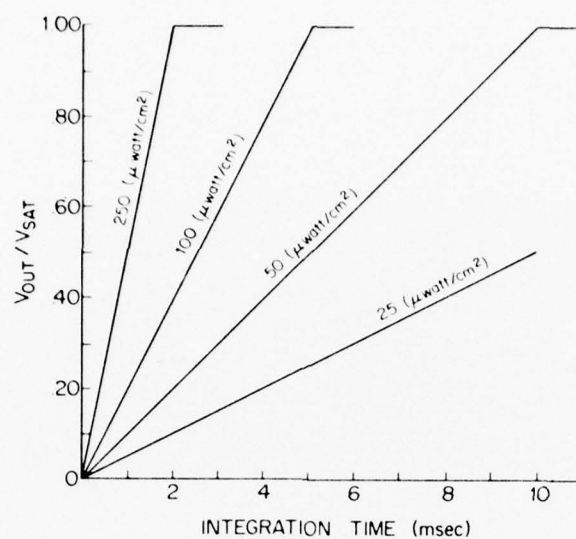
Figure 12.   Relationships Between Light Intensity and Integration Time

(a)   (b)
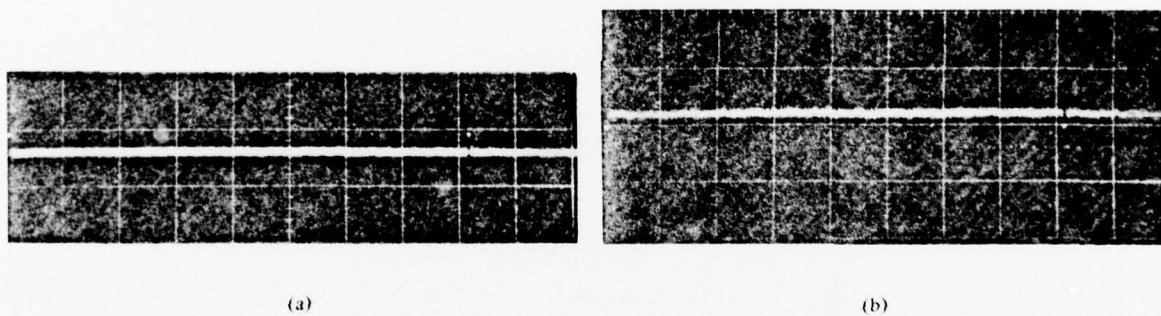
Figure 13.   a) Dark Scan at 1MHz Sample Rate

b) Dark Scan at 100KHz Sample Rate
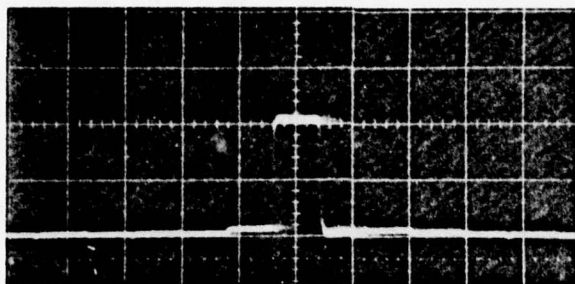
Figure 14.   Boxcar Output with Single Pixel Illuminated
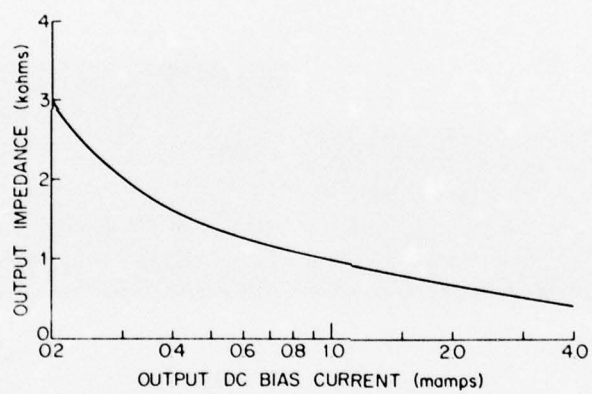
Figure 15.   Output Impedance as a Function of Load Current

LECTURE ET TRAITEMENT ACOUSTIQUES D'IMAGES OPTIQUES (*)

H. Gautier, C. Maerfeld et P. Tournois
Thomson-CSF - Division des Activités Sous-Marines
Chemin des Travails - 06802 Cagnes-sur-Mer (France)

Les ondes acoustiques de surface fournissent une méthode de balayage de senseur optique qui permet une lecture soit directe, soit par traitement de l'image. Dans le domaine de la lecture d'images visibles, l'avantage des techniques acoustiques est qu'elles permettent d'opérer de façon simple un traitement de l'image au moment de la lecture. L'effort consenti à l'étude de tels dispositifs a néanmoins été limité, aussi ces dispositifs sont-ils toujours principalement des objets de laboratoire.

Dans cet article nous rappellerons brièvement la géométrie des dispositifs utilisés ainsi que les principes généraux de fonctionnement. Nous montrerons comment s'effectue la lecture directe d'une image et comment on peut bénéficier du phénomène d'intégration photonique à condition de disposer de senseur de type réseau de diodes intégré. Les systèmes utilisant de tels réseaux seront particulièrement décrits car ce sont ceux qui ont démontré les plus grandes sensibilités.

Nous détaillerons ensuite les possibilités offertes de traitement d'image. En particulier nous montrerons que le signal de sortie peut être la transformée de Fourier ou de Fresnel linéique de l'image. Enfin, nous analyserons les autres capacités non encore démontrées de ce type de systèmes comme la lecture et le traitement bidimensionnels.

1. - PRINCIPE DE LA LECTURE ACOUSTIQUE D'IMAGE.

Tout dispositif permettant la lecture d'image doit remplir deux fonctions : la fonction senseur et la fonction de balayage.

Les senseurs utilisés dans les dispositifs acoustiques sont des mêmes types que ceux des systèmes à tube ou CCD. Ce sont en général des plaquettes de semiconducteur épitaxié ou des mosaïques de diodes diffusées. L'interrogation et le balayage du senseur sont obtenus au moyen d'ondes acoustiques de surface se propageant sur un substrat piézoélectrique, un faible espace d'air étant ménagé entre les deux matériaux (voir Fig. 1a) pour éviter tout effet de charge mécanique des ondes. Du fait de l'effet piézoélectrique, un champ électrique existe qui interagit avec le matériau semiconducteur. Cette interaction peut être simplement décrite au moyen du circuit équivalent de la Fig. 1b. L'onde acoustique est représentée par un générateur de courant I_{ac} proportionnel à l'amplitude de l'onde, et d'admittance interne Y_{ac} ; le semiconducteur d'admittance équivalente Y_s est parcouru par un courant à la fréquence de l'onde acoustique ω et de valeur $j\omega D_{ac}$ où D_{ac} est le champ de déplacement qui existe à la surface du piézoélectrique. L'admittance Y_s dépend en particulier de la densité de porteurs libres n du semiconducteur et de la densité q_s de charges stockées ou piégées à la surface du semiconducteur. On peut donc écrire symboliquement :

$$Y_s = f(n, q_s) \qquad (1)$$

Une des méthodes de balayage acoustique proposées utilise la modulation de Y_s par l'illumination incidente (modulation de n ou de q_s), ce qui perturbe la vitesse de propagation de l'onde et donne lieu à une réflexion de l'onde incidente (QUATE, C.F., 1974). De manière plus générale les techniques de balayage acoustique mettent à profit des effets non linéaires que l'on peut résumer comme suit. Le courant I dans le semiconducteur est de la forme :

$$I = Y_s V + \frac{\partial Y_s}{\partial V} V^2 = I_{LIN} + I_{NL} \qquad (2)$$

et la composante non linéaire peut s'écrire encore :

$$I_{NL} = \frac{\partial Y_s}{\partial V} \cdot \frac{I_{ac}^2}{(Y_s + Y_{ac})^2} = C(n, q_s) I_{ac}^2 \qquad (3)$$

Le terme dérivé traduit l'effet varicap à savoir que la capacité de surface varie avec le potentiel de surface V, c'est-à-dire la charge q_s. Le rendement non linéaire C est ainsi une fonction de n et de q_s.

Le principe de la lecture d'image consiste à faire interagir deux ondes contradirectives de même pulsation ω qui correspondent ainsi à deux ondes de courant $F(t - z/v) \, e^{j(\omega t - kz)}$ et $g(t + z/v) \, e^{j(\omega t + kz)}$ avec $k = \omega/v$ (Voir Fig. 2). Il s'ensuit que :

$$I_{NL} = C \left[F \, e^{j(\omega t - kz)} + G \, e^{j(\omega t + kz)} + C.C. \right]^2 = C \left[A + F G \, e^{j2\omega t} + F G \, e^{j2kz} \right] + C.C. \qquad (4)$$

L'électrode plane déposée sur la face arrière du semiconducteur ne collecte que les courants spatialement uniformes. Le signal de sortie est donc de la forme :

$$H(t) = e^{j2\omega t} \int_{-L/2}^{L/2} F\left(t - \frac{z}{v}\right) G\left(t + \frac{z}{v}\right) C \left[n(z), q_s(z) \right] dz \qquad (5)$$

Si $C = Cste$, $H(t)$ représente le produit de convolution de F et G. Par ailleurs, si on suppose que l'un des signaux est une impulsion brève $F(t) \simeq \delta(t)$, l'autre étant d'amplitude unité $G(t) = 1$, la relation (5) s'écrit :

$$H(t) = e^{j2\omega t} C \left[n(vt), q_s(vt) \right] \qquad (6)$$

Le signal de sortie fournit donc une information sur la valeur de la fonction C à l'abscisse de l'impulsion brève.

Si $I(z)$ est la distribution de l'illumination incidente (Voir Fig. 2), C peut varier locale-
ment :

- soit par effet photoconducteur de modulation de la densité de porteurs libres $n(z) = n_o + aI(z)$

- soit par effet photovoltaïque de modulation du potentiel et de la charge stockée $q_s(z) = q_{so} + b\ I(z)$.

Les ondes de surface balayent donc la densité de charge de surface en lisant la variation du
facteur C ; elles peuvent dont lire tout senseur optique qui transforme intensité lumineuse ou photons en
courant ou charges électriques. Un certain nombre de caractéristiques propres aux techniques acoustiques
de lecture d'image apparaît d'ores et déjà.

a) La vitesse de balayage est égale à celle des ondes acoustiques. Elle est donc élevée. Dans le cas d'un
   substrat piézoélectrique en LiNbO₃, YZ, v = 3500 m/s. Ceci signifie qu'une ligne de 300 points de
   50 μm est lue en 4,3 μs. Nous verrons au § 3 qu'il est néanmoins possible d'augmenter le temps de ligne.

b) Ces dispositifs de lecture fonctionnent généralement dans la gamme VHF. En effet, l'obtention d'une
   bonne résolution Δz demande une large bande passante B, selon la relation $B = v/\Delta z$ (par ex. B = 70 MHz
   pour Δz = 50 μm). Et les transducteurs électro-acoustiques fonctionnent avec une bande relative assez
   faible de l'ordre de 30 à 50 %.

c) Les fonctions "senseur" et "lecteur" sont séparées et elles peuvent donc être optimisées séparément.
   Par exemple, nous avons pu utiliser des cibles de type vidicon faites de matrices de diodes pn. Ces
   cibles présentent en particulier une très bonne isolation entre diodes et un très faible courant inverse
   ce qui nous a permis d'utiliser des techniques d'intégration photonique pour atteindre de très faibles
   seuils de sensibilité à l'éclairement (voir Section 2).

d) La lecture d'image n'est pas destructive contrairement au cas des dispositifs à transfert de charge. Il
   est donc possible de faire plusieurs lectures successives et d'intégrer électroniquement les signaux.
   Ceci rend possible l'obtention directe de transformées globales de l'image. Nous reviendrons sur ce
   point dans la Section 3.

2. - DIFFERENTS MODES DE LECTURE DIRECTE D'IMAGE.

Les dispositifs expérimentaux utilisent en général un substrat en LiNbO₃ qui est transparent
pour le rayonnement visible. Le film d'air nécessaire entre piézoélectrique et semiconducteur est de l'ordre
de 2000 Å pour permettre un bon couplage acousto-électrique aux fréquences acoustiques utilisées qui sont
de l'ordre de 100 MHz. Ce film d'air est généralement obtenu par une attaque ionique du substrat piézo-
électrique qui laisse subsister des plots sur lesquels s'applique le senseur.

Le senseur optique peut être une plaquette de semiconducteur homogène. Mais alors les difficul-
tés liées à la préparation et à la stabilisation de la surface active rendent peu répétables les résultats
concernant la sensibilité de tels dispositifs. La résolution est aussi limitée par la diffusion latérale
de porteurs et la recombinaison de surface.

Une structure plus stable et plus performante est celle dans laquelle le senseur est composé
d'un réseau (ou d'une matrice) de diodes diffusées dans des fenêtres en Silice puis recouvertes d'une
métallisation (Voir Fig. 3). Cette structure a de plus la particularité de pouvoir conserver une charge
négative, charge qui peut être simplement introduite en polarisant dans le sens direct les diodes au moyen
d'une impulsion appliquée extérieurement. Tous les résultats rapportés dans la présente section ont été
obtenus avec des dispositifs de ce type où les diodes diffusées dans du matériau de résistivité 10 Ωcm
portent le code v.m.s. (DEFRANOULD et al, 1976, GAUTIER et al, 1976).

2.1. Lecture de flux photonique (instantané).

Le schéma de principe est celui de la Fig. 2. Sous l'effet de l'éclairement l'apparition d'une
tension photovoltaïque aux bornes des diodes polarise celles-ci en direct. En conséquence, l'amplitude du
signal de convolution H augmente d'une valeur ΔH. La Fig. 4 représente la courbe de sensibilité obtenue
c'est-à-dire la variation de $\Delta H/H$ en fonction du niveau d'illumination. La réponse est loin d'être linéaire
mais il faut remarquer une dynamique d'éclairement de près de 60 dB et un seuil de sensibilité de l'ordre
de 10^{-2} lux en lumière monochromatique de longueur d'onde 0,58 μm.

Dans ce mode de fonctionnement, le signal d'obscurité est élevé et le contraste en sortie
faible. Pour remédier à ces inconvénients, il a été montré (KINO, G.S., 1974) que l'on peut réaliser un
filtrage optique en superposant une grille de pas p sur l'image. Dans ce cas on applique sur les deux
transducteurs d'entrée des signaux de pulsations ω_1 et ω_2 telles que $\omega_1/v - \omega_2/v = \pm 2\pi/p$ (Fig. 5). En
l'absence de lumière le réseau de pas p est inexistant et cette relation de phase ne peut pas être satis-
faite. En présence de lumière, le réseau est projeté sur le semiconducteur et le signal de sortie est
détecté à la pulsation $\omega_1 + \omega_2$. Celles-ci sont choisies de telle manière que le pas p soit suffisamment
petit pour ne pas limiter la résolution.

Pour une modulation du signal H de quelques dB, une dynamique de 30 dB a ainsi été obtenue
(GAUTIER H., 1975), ce qui dans le cas présent, rendrait linéaire la réponse du système.

2.2. Lecture après intégration photonique.

Comme nous l'avons mentionné plus haut, si l'on applique une impulsion de tension extérieure
aux bornes de la structure, avec une polarité telle que les diodes se trouvent passantes, un courant dans
le sens direct circule et charge la capacité des diodes. Celles-ci se retrouvent ensuite polarisées en
inverse et les charges ainsi stockées ne s'éliminent que très lentement au rythme du courant inverse des
diodes. Ce courant est de l'ordre de quelques nA/cm² à température ambiante dans le cas présent et les
charges restent ainsi stockées pendant plus de 10 secondes dans l'osbscurité.

Toute illumination accélère le phénomène de décharge et un temps T après la fin de l'impulsion de charge, la charge stockée aura diminué d'une quantité égale en première approximation au nombre de photons reçu par chaque diode. A cet instant on peut venir relire la charge rémanente en effectuant une expérience de convolution du type de celles décrites précédemment (voir aussi Fig. 6a). Le rendement non linéaire C diminuant avec la charge stockée q_s, le signal obtenu représente une image positive du profil d'illumination.

La Fig. 7 représente les variations du signal de convolution avec le niveau d'éclairement pour divers temps T d'intégration, et dans le cas où la précharge initiale diminue le rendement non linéaire de près de 20 dB. La dynamique d'éclairement est alors limitée à quelques 10 à 20 dB, mais elle est ajustable sur plus de 60 dB en choisissant le temps d'intégration. Des seuils d'éclairement aussi faibles que 10^{-3} µW/cm^2 peuvent être atteints à condition de choisir T ≃ 1 s, ce qui limite d'autant la fréquence de lecture.

Une autre méthode de lecture après intégration photonique est schématisée sur la Fig. 6b. Elle consiste à stocker un réseau de charges par battement de deux ondes contradirectionnelles (ce battement correspond au terme FG e^{j2kz} dans l'Eq. (4)). La charge décroît comme précédemment avec le temps et la charge résiduelle est lue au bout d'un temps T au moyen d'un seule onde brève à la fréquence 2ω ; c'est-à-dire de nombre d'onde 2k adapté à celui du réseau stocké. L'image résultante est alors négative et le réseau de courbes de sensibilité à l'éclairement est donné en Fig. 8. La dynamique d'éclairement est d'environ 20 dB, ajustable comme dans le cas précédent.

Ces modes de balayage par précharge permettent d'obtenir une plus grande dynamique du signal électrique de sortie, ce qui les rend plus faciles d'emploi. La Fig. 9 est un exemple simple d'images obtenues pour trois niveaux d'éclairement différents correspondant à une même énergie lumineuse reçue. La résolution est de l'ordre de 1 mm. Le balayage vertical est obtenu par des moyens mécaniques.

3. - LECTURE PAR TRAITEMENT D'IMAGE.

Dans les modes de balayage par impulsion brève, l'obtention d'une haute résolution nécessite l'utilisation d'impulsions brèves de bande B_c, ce qui diminue le signal de sortie et s. fait au détriment du rapport signal à bruit. Si les signaux sont des codes de longue durée de même larg. c de bande, le signal de sortie représente une transformée de l'image et par un filtrage adapté approprié, on peut opérer la transformation inverse.

Pour démontrer ce principe, considérons le cas de la transformée de Fresnel. Le système est représenté sur la Fig. 10. Deux signaux linéairement modulés en fréquence sont envoyés sur les transducteurs extrêmes du dispositif ; leurs fréquences sont $\omega_1 + \mu_1 t$ et $\omega_2 - \mu_2 t$ respectivement. En présence d'une grille de pas $p = 2\pi v/(\omega_1 - \omega_2)$ on montre que le signal de sortie en présence d'une illumination I(z) est de la forme :

$$H(t) = e^{j(\omega_1 + \omega_2)} \int I(z) \, e^{j\left[\frac{\mu_1 - \mu_2}{2}(t^2 + z^2/v^2) - (\mu_1 - \mu_2) tz/v\right]} \cdot dz \qquad (7)$$

Quand les pentes μ_1 et μ_2 sont différentes, H(t) est la transformée de Fresnel de I(z). Pour un point à l'abscisse z_o, le signal de sortie est une rampe FM de pente $\mu_1 - \mu_2$ et de durée T_c égale à celle des signaux d'entrée et d'excursion de fréquence $B_s = (\mu_1 - \mu_2) T_c/2\pi$. On restitue donc l'image au moyen du filtre dispersif adapté à la pente $\mu_1 - \mu_2$.

On montre alors les trois résultats suivants :

a) La vitesse apparente de balayage est ralentie dans le rapport $(\mu_1 - \mu_2)/(\mu_1 + \mu_2)$ et le temps de ligne T_c peut être beaucoup plus long que le retard acoustique interne.

b) La limite de résolution Δz est fixée par la bande totale de fréquence utilisée $B_c = v/\Delta z$ comme dans le cas du balayage par impulsion brève. Le nombre de points résolus est $N = L/\Delta$.

c) A résolution identique, la technique par transformée d'image permet un gain en rapport signal à bruit d'une valeur $G = T_c B_c$.

Dans l'exemple d'une ligne TV, N = 300, T_c = 60 µs, B_s = 5 MHz on pourrait utiliser un dispositif acoustique avec un retard interne de 6 µs, une bande B_c = 50 MHz. Le gain de traitement serait alors de 35 dB.

Au moyen d'un convoluteur à semiconducteur massif et d'une grille de pas p = 0,175 mm, il a été obtenu les images représentées sur la Fig. 11 (GAUTIER, KINO, SHAW, 1974). Les caractéristiques des signaux d'entrée étaient T_c = 25 µs, B_c = 17,5 MHz et L = 12,5 mm. La résolution est en fait limitée par le pas de la grille et elle vaut environ 0,25 mm.

D'autres transformées sont envisageables. En particulier si les pentes μ_1 et μ_2 sont égales, l'Eq. (7) montre que le signal de sortie est la transformée de Fourier de l'image. Un analyseur de spectre ou un deuxième convoluteur permet de reconstituer l'image. L'image résultante bénéficie, à nouveau, de l'effet d'intégration temporelle et spatiale. Le rapport signal à bruit s'en trouve augmenté et la vitesse de balayage est variable.

Pour ces deux types de transformées, la fréquence dans le signal de la transformée correspond à l'abscisse dans l'image reconstituée. Ainsi toute limitation de la bande du canal de transmission entre le senseur et le décodeur entraînera la perte du pourtour de l'image. Mais on montre que la résolution du reste de l'image sera conservée dans le cas de la transformée de Fourier, ou un peu détériorée dans le cas de la transformée de Fresnel. Les Fig. 11 b,c,d montrent l'effet sur l'image reconstituée d'une limitation artificielle de la largeur de bande du signal transformée. Un filtrage réjecteur des fréquences hautes ou basses du signal transformée de Fresnel, entraîne la perte de l'extrémité droite ou gauche de l'image, sans diminution appréciable de la résolution sur le reste de l'image.

L'emploi des techniques de traitement du signal permet de s'affranchir des problèmes de dynamique : la sensibilité du lecteur n'est plus une fonction du nombre de points résolus. Par ailleurs, ces techniques permettent une plus grande souplesse d'utilisation. Enfin, dans le cas de transmission d'images, il est souvent avantageux de transmettre la transformée de l'image et non son originale, soit pour réduire la bande passante du signal transmis, soit pour réduire l'influence de certains parasites sur le canal de transmission.

4. - EXTENSION AU BALAYAGE BIDIMENSIONNEL.

Nous n'avons présenté jusqu'à maintenant que des dispositifs capables de balayer une ligne d'images. Peu d'études ont en effet été faites à ce jour en vue d'obtenir un balayage acoustique bidimensionnel.

Deux schémas ont été proposés. Le premier (voir Fig. 12a) consiste à juxtaposer les lignes et à les lire séquentiellement. Chaque onde de surface est nécessairement guidée pour améliorer la résolution latérale et la diaphonie entre lignes. Une telle structure peut comporter par exemple autant de transducteurs d'entrée que de lignes, la sortie est commune à toutes les lignes et seules les entrées sont commutées.

Toutes les méthodes de balayage avec ou sant traitement décrites dans les sections 1, 2 et 3, sont applicables ici si l'on suppose une lecture séquentielle des lignes. Il est en outre possible de disposer simultanément de tous les signaux lignes ou de leurs transformées.

Les seuls essais concernant une telle structure ont été faits sur un dispositif comportant trois lignes. Une résolution latérale de 0,4 mm pour une diaphonie de -20 dB a été mesurée (ADAMS et al, 1975).

Nous avons imaginé un autre dispositif qui ne nécessite pas de commutation de lignes. Il est basé sur l'interaction non linéaire de deux ondes $(\omega_1, \vec{k}_1)$ et $(\omega_2, \vec{k}_2)$ se propageant suivant des directions différentes. Le schéma de principe est représenté sur la Fig. 12b. Deux ondes planes issues de deux transducteurs T_1 et T_2, se croisent en un point P. De la même façon que dans le cas colinéaire, si le milieu est non linéaire et photosensible, l'interaction des deux signaux donnera naissance à un signal dont l'amplitude sera fonction de l'éclairement du point P. Celui-ci se déplace à vitesse constante sur une droite qui est le lieu des points d'intersection des deux plans d'onde. Pour balayer toute une surface, il suffit de faire varier à chaque récurrence le temps de retard entre les deux impulsions d'entrée. La fréquence Ω et le vecteur d'onde $\vec{K}$ du signal de sortie sont donnés par :

$$\Omega = \omega_1 \pm \omega_2 \quad ; \quad \vec{K} = \vec{k}_1 \pm \vec{k}_2 \qquad (8)$$

La vitesse de phase Ω/K du signal d'interaction ne correspondant pas à un mode propre de propagation, il est nécessaire de recueillir le signal de sortie en réalisant un accord de phase au moyen d'un réseau d'électrodes de pas $p = 2\pi/K$. Un réseau de diodes tel que celui décrit précédemment pourrait être utilisé à cette fin. Il est aussi possible de stocker le réseau de pas p sous forme d'un réseau de charge.

Enfin, il faut noter qu'il est possible de lire la transformée bidimensionnelle de Fourier de l'image en opérant un balayage ligne à ligne dans l'espace transformé. Il est alors nécessaire de disposer de plus de 2 transducteurs et de signaux modulés en fréquence. La reconstruction de l'image est alors envisageable au moyen de dispositifs analogues pouvant mémoriser sur une surface le signal transformé. Ces principes n'ont pas encore été vérifiés expérimentalement mais ils peuvent ouvrir la voie à de nouvelles techniques de reconnaissance de forme.

En conclusion, nous avons essayé de comparer dans le Tableau I les principales caractéristiques des systèmes de lecture d'image utilisant des dispositifs à transfert de charge (CCD) ou à injection de charge (CID) et de ceux qui employent des ondes acoustiques de surface OAS.

- TABLE I -

Principales caractéristiques des dispositifs de lecture d'images optiques employant des CCD/CID ou des OAS

Dispositifs CCD ou CID		Dispositifs à OAS
Bande de base		Haute fréquence (VHF)
Echantillonné		Analogique
Monolithique		Milieux séparés. (Une version intégrée existe)
Lecture CCD	Lecture CID	Lecture acoustique
----------------	----------------	--------------------------
Séquentielle ou parallèle des lignes	Aléatoire des points	. Séquentielle ou parallèle des lignes
Destructive	Non destructive	. Non destructive
		. Avec traitement d'image
Résolution obtenue ≳ 50 µm à 100 µm		
Seuil de sensibilité ≳ 10^{-3} µW/cm^2		
Dynamique d'éclairement ≳ 50 dB		

Dans les systèmes à CCD/CID les fonctions senseur et lecteur sont intégrées dans un même substrat semiconducteur, ce qui rend la structure monolithique. Les systèmes de lecture acoustique décrits dans cet article présentent un film d'air et nécessitent un plaquage mécanique du senseur sur le substrat piézoélectrique. Des structures stables et homogènes ont été ainsi réalisées. Parallèlement la recherche de solutions monolithiques se poursuit. Il faut citer en particulier les structures intégrées dans lesquelles une couche piézoélectrique mince de ZnO est déposée sur le senseur semiconducteur en Silicium (KHURI-YAKUB, 1974). Des transducteurs interdigités sont évaporés sur la surface du matériau piézoélectrique qui permettent de générer les ondes acoustiques nécessaires au balayage. Des lecteurs optiques unidimensionnels ont été construits qui ont une largeur de bande de 20 MHz correspondant à une résolution de quelques centaines de μm.

Les caractéristiques de sensibilité à l'éclairement et de résolution des systèmes à CCD/CID ou OAS sont similaires car les senseurs utilisés sont du même type. Les valeurs obtenues sont de l'ordre de 10^{-2} lux et 100 μm respectivement.

Enfin, la lecture acoustique est essentiellement non destructive et les dispositifs comportent en général une électrode intégratrice : ceci permet de traiter "in-situ" l'information optique. Cette option n'a pas pas d'équivalent dans les autres techniques. Elle est particulièrement intéressante lorsqu'il s'agit de transmettre l'information ; et elle doit s'avérer très utile dans les systèmes d'imagerie bidimensionnelle, permettant ainsi d'obtenir des transformées bidimensionnelles d'images (spectre, corrélation ...).

-:-:-:-:-

REFERENCES

[1] C.F. QUATE, "Optical Image Scanning with Acoustic Surface Waves"
I.E.E.E. Trans. Sonics and Ultrasonics, SU-21, 283-288 (1974)

[2] Ph. DEFRANOULD, H. GAUTIER, C. MAERFELD "High Sensitivity Acoustically Scanned Optical Imaging Device using Charge Storage Effect"
Appl. Phys. Lett., 29, 2 (1976)

[3] G.S. KINO "Acoustical Scanning of Optical and Acoustic Images", 23rd Microwave Research Inst. Symp. on Optical and Acoustical Micro-Electronics, New-York, April 16-18 (1974)

[4] H. GAUTIER, Ph. D. Dissertation, M.L. Report n° 2448, Stanford University, USA (1975)

[5] H. GAUTIER, G.S. KINO, H.J. SHAW "Acoustic Transform Techniques Applied to Optical Imaging"
1974 I.E.E.E. Ultrasonics Symp. Proc. 99-103 (1974)

[6] P.L. ADAMS, H.J. SHAW, D.K. WINSLOW, L.T. ZITELLI "Waveguides for Transverse Resolution in Surface Acoustic Wave Optical Sensors", 1975 I.E.E.E. Ultrasonics Symp. Proc. 153-155 (1975)

[7] B.T. KHURI-YAKUB, G.S. KINO "A Monolithic Zinc-Oxide on Silicon Convolver"
Appl. Phys. Lett., 25, 188-190 (1974)

-:-:-:-:-:-:-:-:-:-:-:-

DISCUSSION

J.J.Stapleton

What benefit does your technique have over the modulation of the vidicon ($Sb_2 S_3$) thermionic cathode with the chirp or with feedback from the photocathode?

Author's Reply

It is non-destructive. We are concerned with the transform itself more than the image.
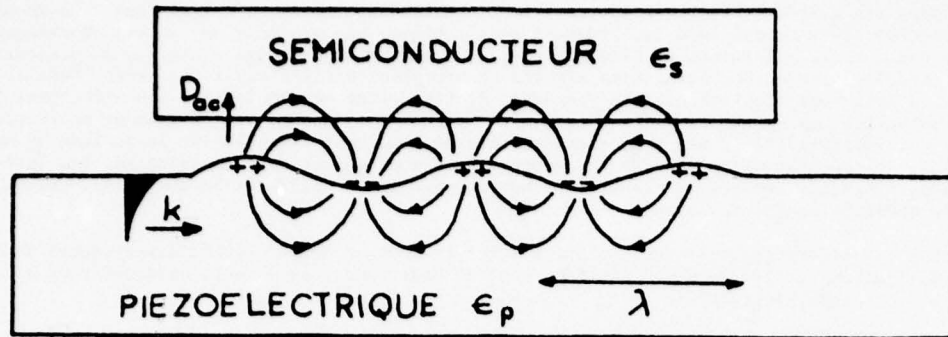
(a)
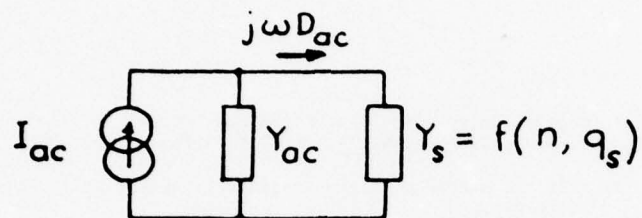
(b)

Fig. 1 - a) Principe d'interactions acousto-électriques entre ondes de surface et semiconducteur
b) Circuit équivalent linéaire.
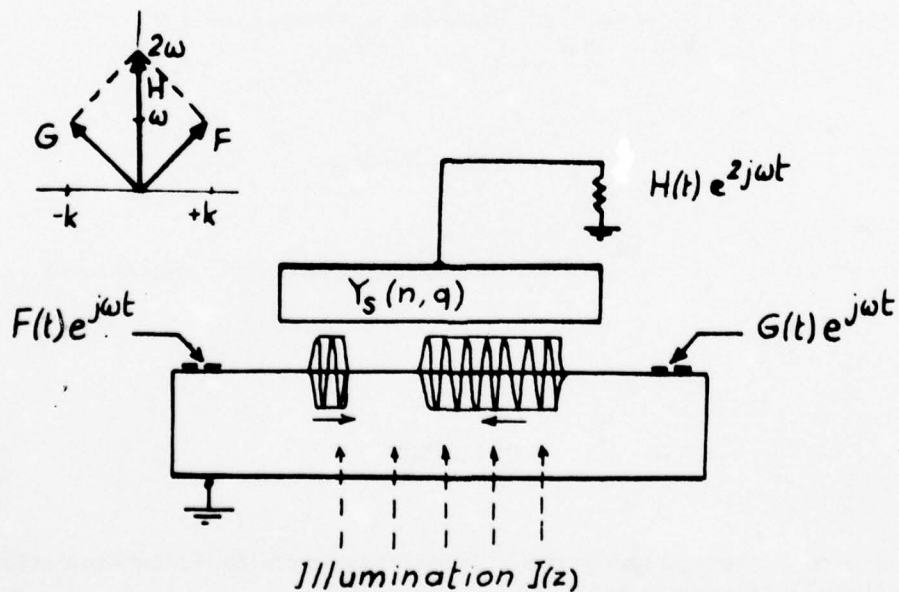
Fig. 2 - Principe de la lecture acoustique d'images optiques.
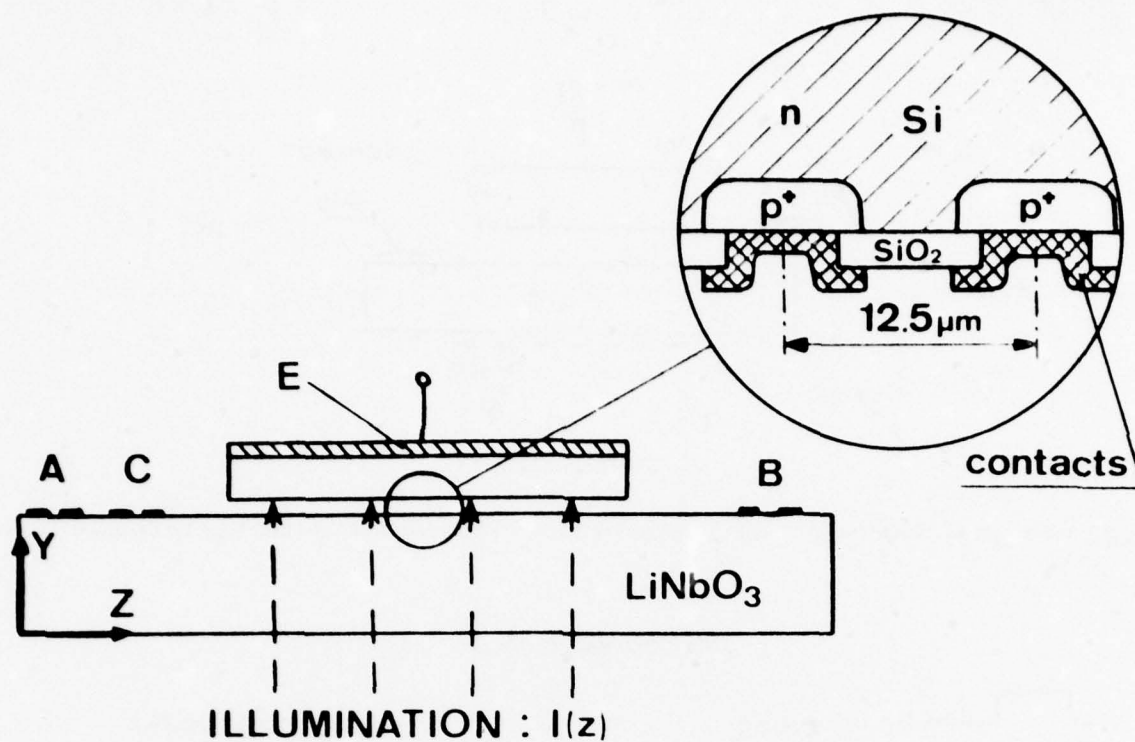
Fig. 3 - Dispositif expérimental utilisant une cible vidicon v.m.s. comme senseur optique.
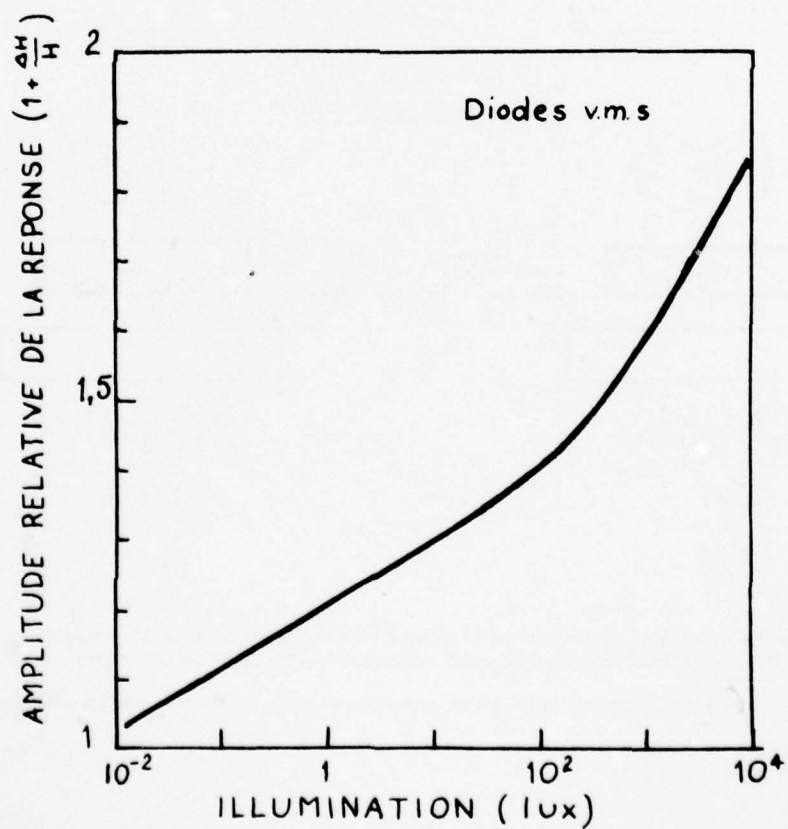
Fig. 4 - Sensibilité relative à l'éclairement. L'illumination module le rendement du système utilisé en convoluteur simple.
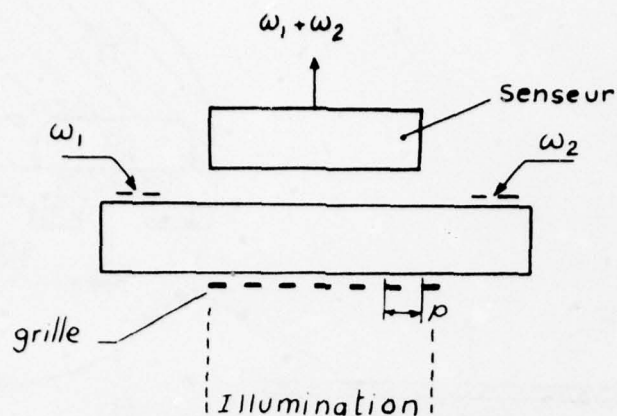
Fig. 5 - Méthode de réduction du signal d'obscurité par introduction d'une grille dans le faisceau optique
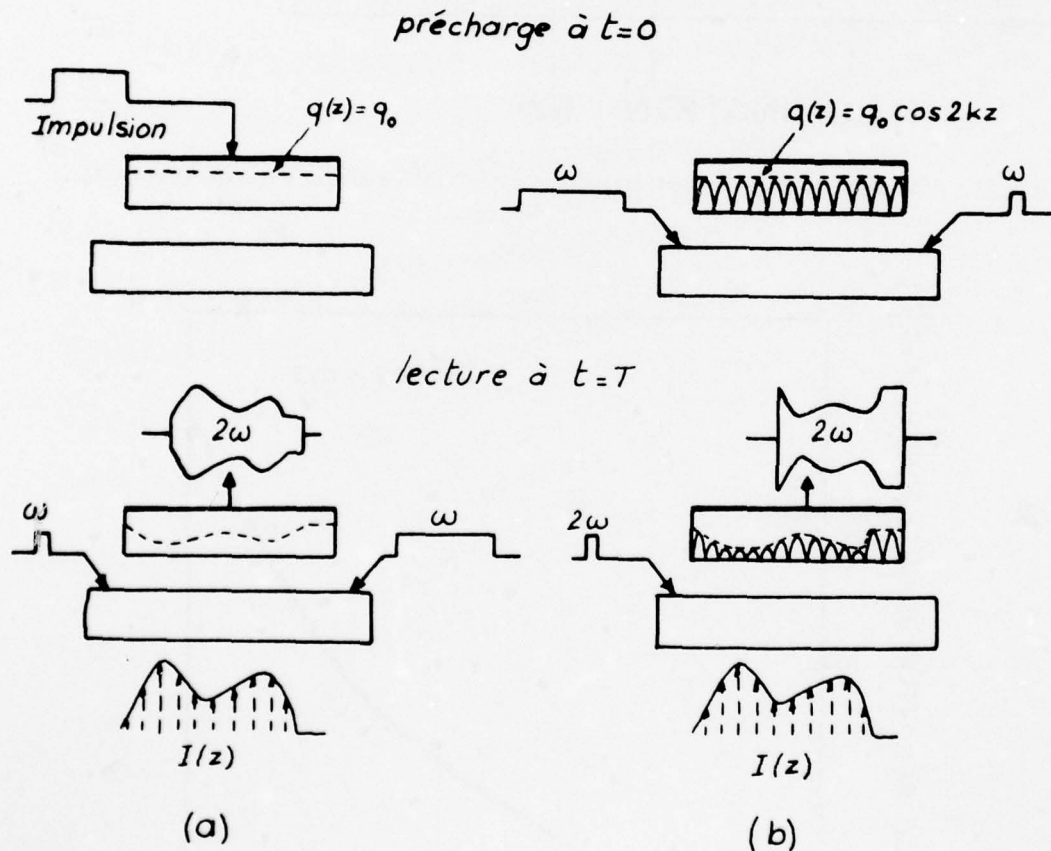
Fig. 6 - Méthode d'imagerie bénéficiant de l'intégration de photons dans le senseur :
  a) Précharge uniforme des diodes puis lecture par interaction non linéaire de deux ondes à la même fréquence
  b) Précharge spatiallement périodique des diodes puis lecture au moyen d'une onde acoustique.
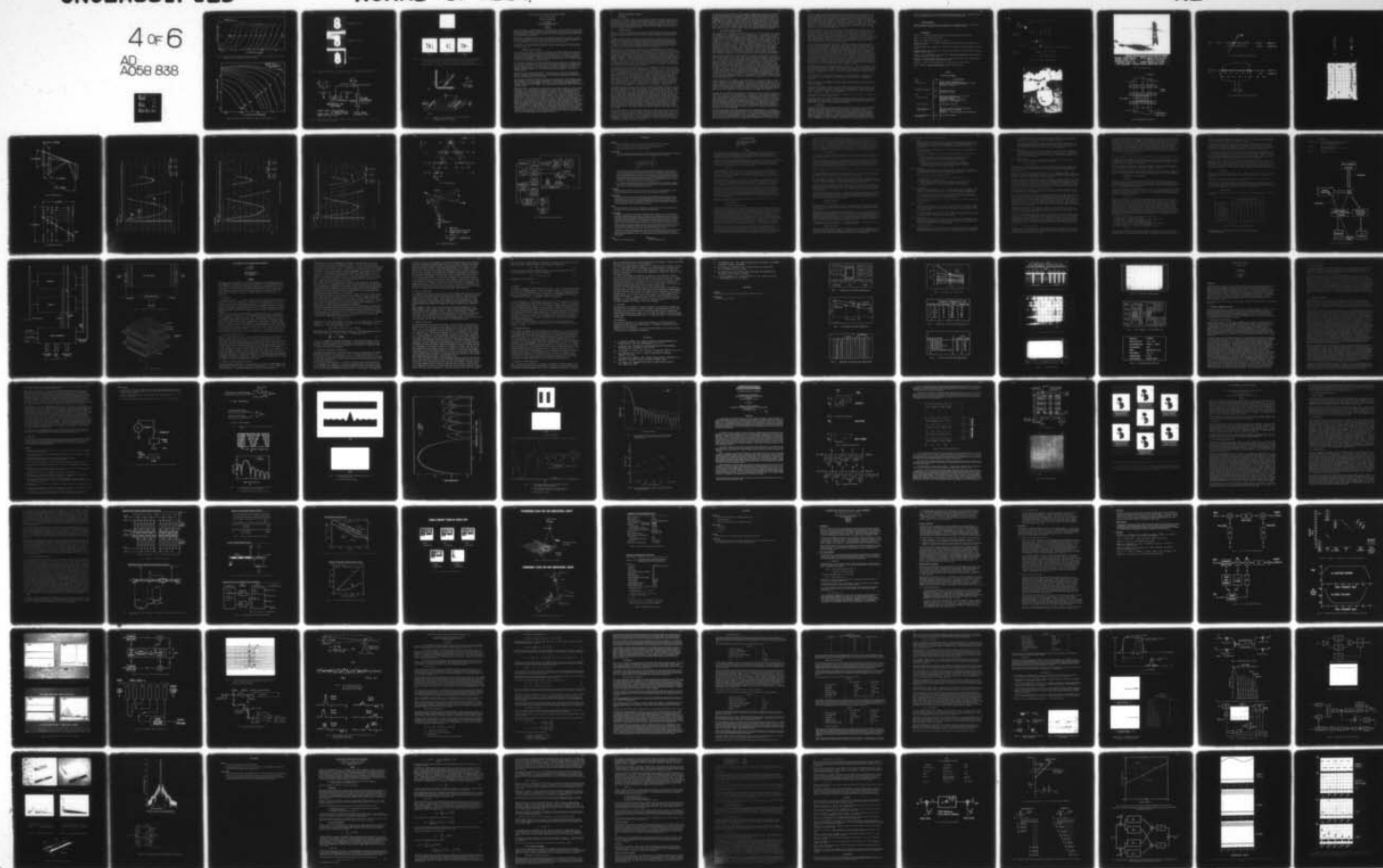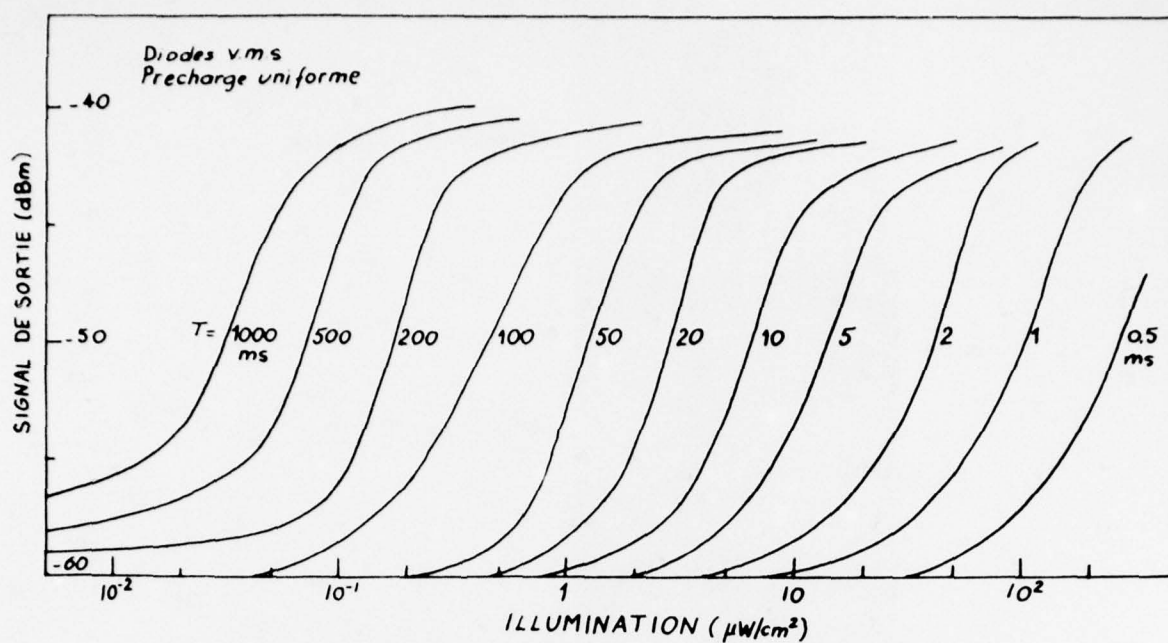
Fig. 7 - Sensibilité à l'éclairement dans le mode de lecture de la Fig. 6a. Le paramètre T est le temps séparant l'instant de précharge de celui de la lecture.

Fig. 8 - Sensibilité à l'éclairement dans le mode de lecture de la Fig. 6b.

$120\,\mu W/cm^2,\ T=120\,\mu s$

$45\,\mu W/cm^2,\ T=380\,\mu s$

$15\,\mu W/cm^2,\ T=1\,ms$

$\vdash\ 1\,\mu s$
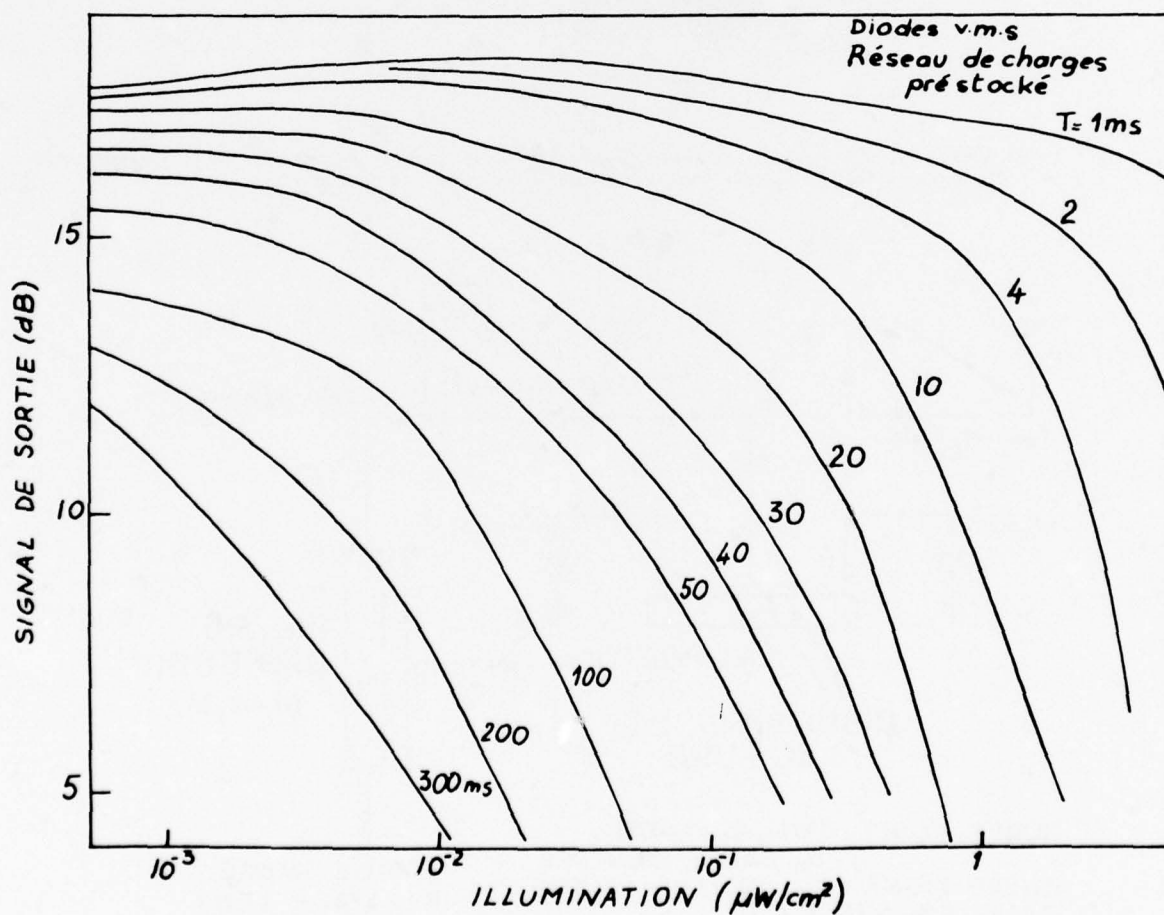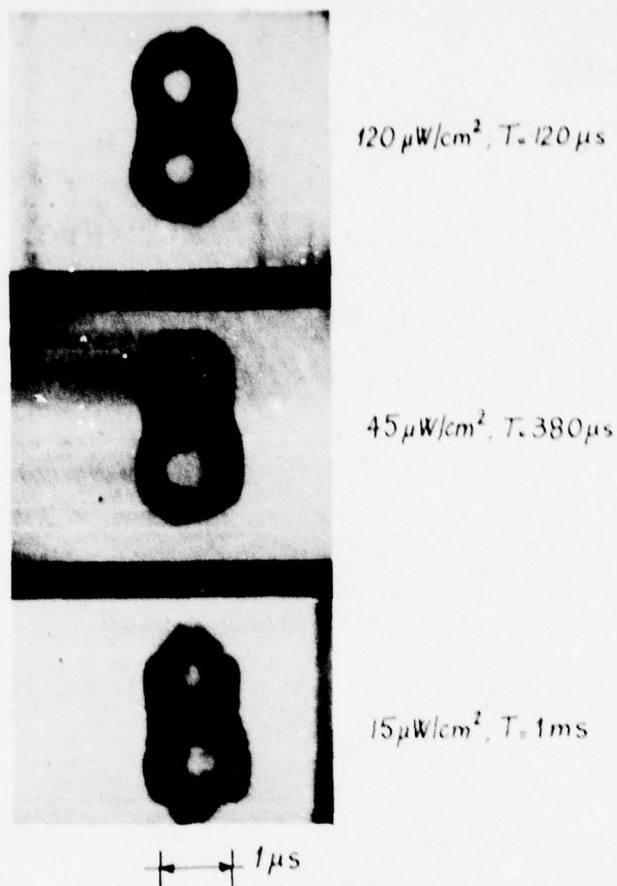
Fig. 9 - Image d'un 8 obtenue avec le système à diodes v.m.s. et la méthode par précharge spatiallement périodique (Fig. 6b). Les trois images sont obtenues pour une même énergie lumineuse.

ILLUMINATION

DISPOSITIF DE LECTURE

RETARD T_C

FILTRE DISPERSIF $(\mu_1 - \mu_2)t$

SORTIE IMAGE BALAYAGE LENT

DUREE POINT : $2\pi/(\mu_1 + \mu_2)\,T_C = 1/B_C$
B_C BANDE D'ENTREE
DUREE LIGNE : T_C RETARD FILTRE
GAIN EN S/B = $T_C\,B_C$
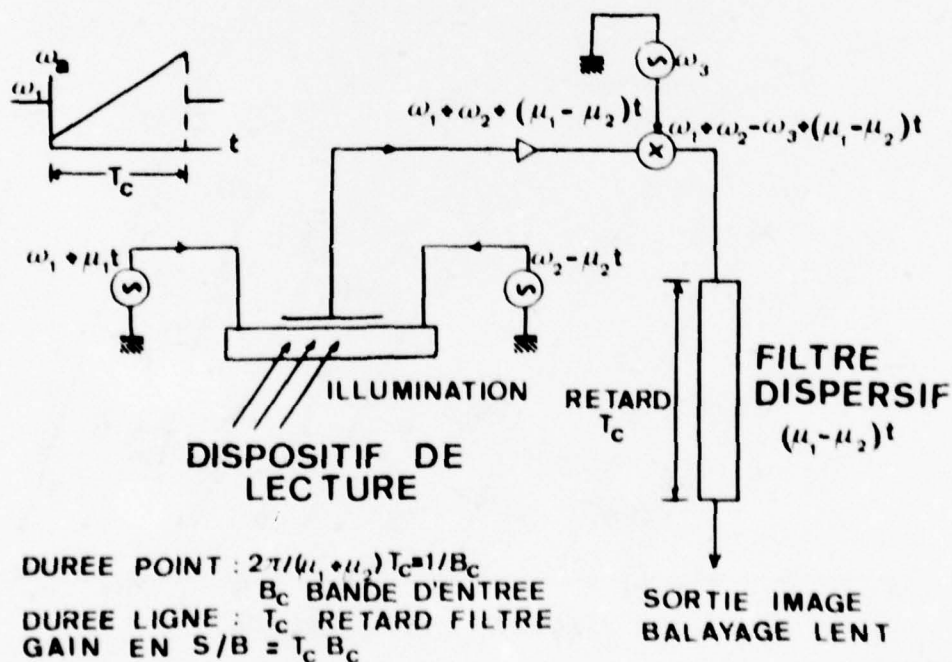
Fig. 10 - Schéma du dispositif de lecture d'une image optique par transformée de Fresnel et reconstitution dans un filtre dispersif adapté.

(a)

(b)　　　　　(c)　　　　　(d)
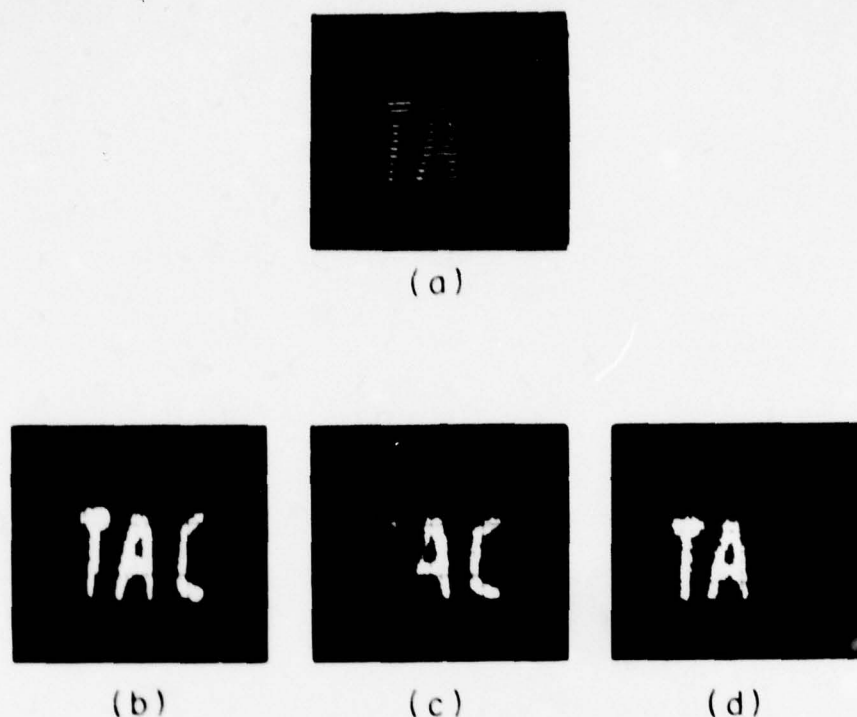
Fig. 11 - a) Image prise avec le système par transformée de Fresnel (Ref. 4). Les rampes de fréquences sont centrées autour de 110 et 130 MHz, leur durée est 25 μs et leur excursion de fréquence 9,8 et 7,6 MHz respectivement. Une grille de 0,175 mm de périodicité est utilisée.

b,c,d) Effet d'une limitation de bande du signal transformé sur l'image reconstituée.
b) Aucune limitation. c) Suppression des basses fréquences. d) Suppression des hautes fréquences.
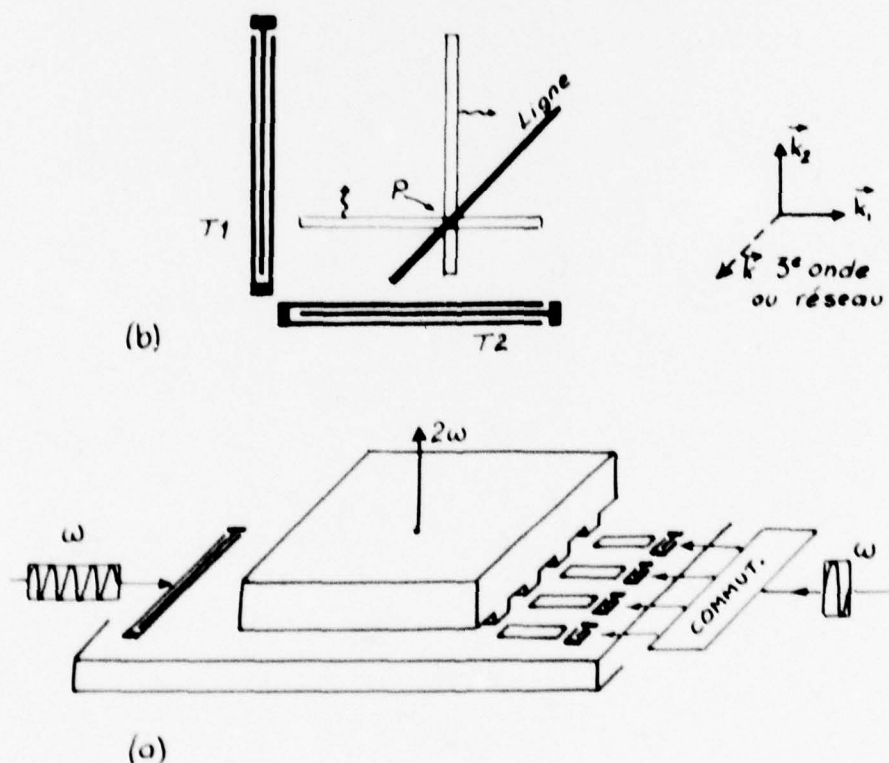
Fig. 12 - Schémas de principe du balayage acoustique bidimensionnel
a) Balayage par ligne commutables
b) Balayage par ondes non colinéaires.

APPLICATIONS OF A CHARGE COUPLED DEVICE SENSOR

FOR NAP-OF-THE-EARTH

HELICOPTER OPERATIONS

A. KLEIDER
US Army Electronics Command
Fort Monmouth, NJ

SUMMARY

The unique geometric characteristic of a wire obstacle has been conceptually integrated with the discrete elemental structure of a Charge Couple Device (CCD). This quantized detector structure provides a means for formulating a Wire Obstacle Warning System (WOWS) utilizing advances in technology to arrive at a low cost system for use by helicopters flying Nap-of-the-Earth (NOE).

WOWS employs a CCD detector array and other monolithic semiconductor devices to achieve small size, lightweight, and low cost. The video information is logically processed to provide a symbolic display of the range and location of the wire obstacle relative to the aircraft heading as well as an audio/visual alarm for the pilot.

The concept of single site activation (SSA) provides a means of achieving wire pattern recognition which removes the requirement of having a human observer in the detection/recognition loop.

1. GENERAL PROBLEM & HISTORY

The problem of wire strike incidents as related to helicopter operations has been evident since the earliest days. The delicate stability of these aircraft leave them particularly vunerable to contacts with even the smallest wires. In addition, in many instances loose wires become enmeshed in the rotor and mast, thereby interferring with the flight control systems. Obviously, contact with larger wires such as power lines, are usually, catastrophic. In fact a single 6cm diameter power line in VietNam accounted for no less than fourteen separate crashes.

The statistical history of helicopter wire strikes reflects the costs in human and fiscal resources. Even neglecting the exigency of war time operations, wire strikes continue to plague the Army and civilian aviation community to the extent of approximately four incidents per month. These statistical facts must be weighed in the light of existing operational doctrine. These doctrines primarily reflect (1) the tactical employment of such aircraft, and (2) the survivability of helicopters in the age of "smart" weapons.

The tactical dictates related to simple survival include (1) Nap-of-the-Earth (NOE) operation, and (2) nighttime operation. A third dictate includes an all weather capability. However, from a practical point of view this must be modified to weather conditions less than ideal but not zero-zero.

The meanings of these dictates and their precise definitions are worth including here so that the extent of the wire obstacle problem can be appreciated. The term NOE includes flight regimes at altitudes under twenty meters. In some instances the skid/terrain clearance may be one to two meters. The term nighttime includes twilight (3×10^{-1} foot candles) to dawn (40 foot candles) with illumination levels as low as 6×10^{-5} foot candles (moonless night).

These definitions, combined with the operational specifications relative to the aircraft forward velocity, lift capabilities, and pilot reaction time, delineates the nature of the problem. In the case of wire obstacles the physical size of the objects and the distances at which detection and recognition must occur if an avoidance maneuver is to be accomplished not only makes the problem a difficult one but also limits the techniques which can be applied to achieve the desired ends. Practical dictates for the parameters of concern provide for wires as small as 3 millimeters at ranges of 300 meters to 500 meters. Thus the targets encompass a field of view (FOV) in one dimension of from 6 to 10 micro-radians. This number is an order of magnitude below the resolution capability of the human eye for high contrast targets. Pilots equipped with night vision devices can clearly discern most objects in quarter moonlight situations to permit the nighttime equivalent of Visual Flight Rules (VFR) operation. With specially trained pilots the adapted but unaided human eye is sufficient to permit night flight in familiar areas. However none of these techniques is remotely possible where wires are likely to be encountered. It is imperative that a reliable means of detecting, recognizing, and avoiding (DRA) wire and wire like objects be developed. Such a device is an absolute necessity if the Army is to have the capability of using helicopters in a war time situation and equally necessary if real time training of pilots for such NOE/nighttime operations is to be accomplished with any margin of safety.

2.   EMERGENT TECHNIQUES & CONCEPTS

2.1   BACKGROUND

The problem of a suitable solution to reliable wire detection has been investigated for the past ten years. Figure 1 shows a techniques tree of successful and unsuccessful methods for wire obstacle avoidance systems which have been explored at the US Army Avionics Laboratory, Fort Monmouth, NJ. Table 1 expands the detailed coverage of these techniques. These techniques merely serve as a historical backdrop for the focus of this presentation which is the Wire Obstacle Warning System (WOWS) whose principal component is a Charge Coupled Device (CCD).

2.2.   THE WOWS CONCEPT

The use of a CCD in the WOWS constitutes a conceptual breakthrough in the wire DRA environment, and holds the promise of providing a solution to the wire strike problem with the desirable properties of being lighter, smaller, cheaper, and above all, reliable.

The WOWS program represents a marriage of matching characteristics relating to the appropriate description of the targets and the physics of the detection/recognition device. This compatibility of detector and target is expressed by the principal goal of WOWS, i.e. to detect and discriminate wires from everything else within the field of view of the device. This is possible because a wire has a visual signature which is unique. In a sense the very physical essence which creates the problem of wire detection and recognition also provides the key to its solution.

While the world of our everyday experience is three dimensional and we are familiar with conventional two dimensional representations, there are few quasi-one dimensional objects which exist in our experience. The smallness of the diameter of the wire strains the acuity of the human eye and other optical systems, while the extended nature of its length provides a high degree of regularity which is useful and important for WOWS. Fundamentally WOWS is an automated recognition system and does not constitute an imaging device. The unique signature of the wire as a one-dimensional object combined with the geometric regularity in the extended dimension provides an ideal match to the linear CCD 1xM array. Where the value of M in the array is large, such as the Fairchild CCD 121 (M=1728), it is possible to get very high resolution. By orienting the high resolution axis orthogonal, or at a large angle to the linear dimension of the wire good detection cross-section can be obtained. Simultaneously a coarser scan orthognal to the diameter will accurately exhibit the geometric regularity.

Physical insight into the processes which govern the WOWS were derived from the results of the Gated, Low Light Level TV (GV^{3}TV) program. [Kleider, A., '75]. The success of the GL^{3}TV technique was directly attributable to the sidelobe structure of the Si target in the Westinghouse WX-31677 Intensified Silicon Target (ISIT) tube and a VL-31793 Camera tube. In systems employing other intensifier/read-beam systems their limited MTF capabilities precluded the resolution of wire objects. This lack of resolution capability combined with the low Q characteristic of the point spread function for small sized objects in an ISEC type tube rendered them incapable of detecting wires. However, it was initially discovered experimentally that wires as small as 3mm and as far away as 3Km could be discerned for an ISIT based system with a pulse illuminator and a narrow-gated intensifier camera. In this test a 410mm, T2.0 lens with a 5.6 degree FOV was used. Figure 2 shows the camera mounted on a UH-1 helicopter and figure 3 shows the results obtained at a range of 2.25 kilometers in a ground test. The wires shown are 3cm diameter and 6mm diameter as illuminated by a 0.6 watt GaAs, 15KHz system.

The important feature of these results was the apparent and reliable display of objects beyond the resolving capability of the system. The hypothesis which evolved to explain this result attributed the observed performance to the discrete nature of the Si target in the ISIT [Kleider, A., '74]. The individual Si diodes which constituted the information storage media in the image tube were approximately 5 microns on each side. In practice the electron read beam encompassed more than one diode and the resultant video was an average over those diodes within the read beam diameter. However the discrete nature of the target effectively quantized the low Q point spread function which characterized the wire while the extended dimension produced sufficient contiguous charged elements to yield a suitable video output. Essentially it was postulated that while the object may be imaged on less than a whole receptor, the receptor can take on a charge proportional to the total integrated intensity and respond accordingly. The "image" is not resolved but its presence is discerned. Thus for a wire whose diameter is twice that of a similar wire, but where neither wire completely fills the detector element, the resultant "image" can not be measured to provide information concerning the size of each of the wires, ie, the larger wire may be twice as bright but the "image" size will be the same.

The apparent success of the GL^{3}TV system with the ISIT tube was clouded by several negative factors. These include the probable high cost of such a device and the need to provide a dedicated crew member to monitor the display while operating NOE. This latter factor is by far the more serious and limiting in that the variables in human preception of displays are unacceptable for the establishment of meaningful Standard Operating Procedure for nighttime NOE operations. However, the results of the GL^{3}TV program and the analysis of those results have provided the key to the concept which will provide a predictable, reliable, and low cost system.

The recognition of the importance of the unique one dimensional signature of wires and the quantum-like detector for a system designed to detect wires was significantly enhanced by the technological development of Charge Coupled semiconductor Devices (CCD). Whereas the size and cost of the CCD as compared to an ISIT is an important reason for considering CCD's in a wire detection system, the most important reason for their use is the nature of the device. Unlike the ISIT the individual elements of the CCD can be probed and their information content used to detect and recognize wires. The physical size of the CCD elements, as shown in figure 4, are comparable to the elemental diodes in ISIT. Since wires of interest and concern have angular intercepts in the range of 0.1 mr to 0.01 mr, their image in the focal plane of the CCD constitutes less than one pixel. The ability to interogate each pixel in the CCD combined with the one dimensional signature of the wire can be merged to yield a suitable recognition algorithm.

2.3   WIRE OBSTACLE WARNING SYSTEM

Details pertaining to the physics and electronics of the CCD can be found in the literature [Amelio, G.F., '74; Barbe, D.F., '75,'76; Jespers, P.G., '76; Sequin, '75], however it is appropriate to comment upon a unique feature of CCD's. Whereas previous semiconductor devices such as transistors result in modulated voltages or currents the CCD is effectively a device which manipulates information bytes. These "bytes" can be two leveled, ie., "0" or "1" in standard digital format or they can be multileveled such as a shades of grey scale to produce an image. Simplistically a CCD possesses three attributes; (1) An input; (2) A charge coupling mechanism, and (3) An output. The input stores quantized bytes of "scene" information which are transferred in parallel upon command to the transfer register. The charge coupling mechanism moves these bytes sequentially through the register to the output which constitutes the video signal. The combination of wire signature, "byte" construction of the CCD, and the sequential incremental processing of the fundamental video chain yields a simple and powerful means of detecting and recognizing wires. The device is not an imaging one in the common sense of the word. It is a pattern recognition processer which is combined with other electronic subsystems to form an automatic Wire Obstacle Warning System (WOWS)

Assuming that only the wire is within the FOV of the array, (See Figure 5) and there is no background present, if the video output is digitized, only element 18 would be nonzero. For an object such as a tree, (See Figure 6) elements 22 through 32 would be nonzero. Since the intent is to discriminate wires from all other objects in the world the concept of Single Site Activation (SSA) has been implemented. SSA is defined as the digital algorithm of [010]. Orthogonal to its diameter the wire is represented by a continuous extended target whose geometry is very nearly linear (sectionally). This provides the necessary relationship between SSA and the "scene" scanning mechanism to describe the wire to the decision logic. SSA outputs on a Fairchild 1x1024 linear array are shown in figure 7. The lower scope trace is the video output covering approximately 12 pixels. The SSA for two one mil wires is clearly evident. The upper trace mirrors the lower trace except that the video has been thresholded. The data shown represents a single scan line within the projected FOV which the WOWS system would examine. Assuming the total FOV is $\pm$ 5° in elevation and $\pm$ 15° in azimuth (corresponding to an area 52.3 Meters by 157 Meters at a distance of 300 meters) the system would divide the space scanned into three hundred slices with an instantaneous FOV of 175×10^{-3} radians in elevation by 10^{-4} radian in azimuth. Each slice would then be separated by 1.75×10^{-3} radians. This is shown in figure 8.

Figure 9 shows a matrix N elements long by M scans wide. To each SSA detection found in a single frame there is assigned a number pair which can be stored in memory and displayed on command. The extended geometry of the wire is then used as a means of identifying the SSA pattern as a wire obstacle. Several modes of operation are possible. The simplest device would display on command at the end of each frame the SSA. An observer could readily discern a wire presence even where a moderate random background noise was present. A typical display of this type is shown in figure 10 for several backgrounds. These results were produced by a computer simulation of a WOWS type of system.

A second and more sophisticated approach treats each SSA found in a scan as either the beginning or continuation of wire detection. Depending upon the separation between scans a gate, G_1 on figure 11 is selected about X_i on scan Y_i+1. This section of the scan line is then emphasized and searched for SSA. Assuming SSA are found on X_i-3 and X_i+7 a second gate, G_2, on scan Y_i+2 is set and this area searched on the Y_i+2 scan. Each time a correlation is found within the gate set, the next one is narrowed. Allowing for deviations from the predicted line shape and drop outs, the number of correlations required to establish the presence of a wire will reflect the system false alarm rate. Initially the probability of detection is set high and since the expected signal to noise ratio (S/N) per element is low (implying a high false alarm rate) the statistical products of the Poisson distribution rapidly coverages to a system false alarm rate acceptable for real time systems.

The practical system operational requirements imply a high probability of detections, which is reflected by a low threshold setting in the signal digitizer, and a probability of false alarms sufficiently small such that in the time associated with NOE operations none or at most one false alarm will occur. The factors which bear upon these values are (1) the S/N, (2) the background, (3) the wire shape, (4) the number of continuities required in the decision algorithm, and (5) the detection array configuration. The desired system operating values are a probability of detection, P_d of 0.98 and a false

alarm rate, P_{fa} of 10^{-6} which would yield one false alarm per hour of operation. The design of a WOWS includes a trade off analysis of these factors. The S/N and the continuity requirements as they influence the system logic is of particular importance. In the former the S/N is directly tied to the laser illuminator output power and both cost and weight are proportional to this characteristic. In the latter, the physical memory requirements of the computor sub-system are intimately related to the convergence requirements. Here again system size, weight, and costs are at stake. Our early studies have yielded peak power requirements of less than 3KW with an average power of approximately 0.1W. Continuity requirements range from 5 to 7 in low noise situations to a maximum of 13 in a high background case. These values have been verified in our WOWS simulation program [Fairchild Camera Corp., '76].

As shown in figure 1, passive systems have not been successful as wire detectors. The WOWS concept thus includes a laser illuminator. For convertness and potential low input power operation a GaAs diode laser has been selected. With its principal output at 0.9 microns, which is beyond the responsivity of the unaided eye, the GaAs is well matched to the response of the CCD Si base material. Other laser illuminators are potentially useful for this mode of operation. YAG lasers operating at 1.06u can provide 10KW -20KW peak power outputs with good efficiency. However,at 1.06u the Si response is down by approximately 3db and slight changes in output wavelength as the YAG heats up can result in drastic changes in detectivity since the response curve for Si falls rapidly in the 1.0u to 1.1u region of the spectrum.

A short pulse illuminator synchronized with a range gated detector provides a system free of foreground atmospheric scattering. Similarly this gated operation permits us to view a restricted volume of space and thereby reduce the "scene" background which would tend to limit the S/N by imposing a high threshold value for the digitizer. However the principal benefit is that the gated operation provides range data. In operation the system will have an adjustable, but fixed, range-gate setting. Azimuthal and elevational scans are part of the WOWS however the longitudinal scan is performed by the aircraft itself as shown in figure 12. Since the WOWS logic is capable of recognizing a wire after n continuities which fit the selected algorithm, the total FOV which is covered by M scans is divided into (M/n) decision elements. The corresponding decision FOV, θ_d, is $\theta_t/(M/n)$. The distance the aircraft moves per frame is v.ts so that in one decision time the range "blur" is $(v.ts)/(M/n)$. Typically the aircraft velocity in NOE is 30M/sec. Estimate of the decision time for WOWS is 50×10^{-3} sec which yields a range blur of 1.5 meters. When an initial wire warning is recognized by the system the range indication is decremented by the on-board computer in accordance with the measured ground speed until such time that the pilot releases the system.

The WOWS configuration is shown schematically in figure 13. The precise design is still being evaluated, however the features depicted are accurate. The CCD selected for this flyable exploratory model of WOWS is the Fairchild CCD 131 which is a 1024 linear array. The CCD will be integrally affixed to an 40 mm ITT microchannel-plate amplifier (MCP) coupled to a fiber optic minifier. The MCP has external provisions for gating the intensifier on and off.

The illuminator is a GaAs pulsed array with a peak power capability of 3KW and a variable pulse width in the range of 50 nanoseconds to 100 nanoseconds. Illumination uniformity is achieved with an optical integrator interfacing with the illuminator field optic.

The logic, memory, and computer are microprocessor systems similar to the Intel 8080. Where ever possible LSI components will be employed to minimize size, weight, cost, and power requirements.

The scanning mechanism consists of pair of stepped rotary tables driven by precise scan generators. No attempt is made to provide a stabilized platform, however scan to scan results are sensed and computer reconstructed to accommodate for aircraft attitudinal changes. This data is derived from an AN/ASN-118 on-board sensor. This technique is roughly equivalent to image motion stabilization.

The total FOV of the system is ± 6 degrees in elevation and ± 20 degrees in azimuth centered about the aircraft flight vector.

The display can be a conventional CRT with an appropriate symbol generator. However the display does not require high resolution therefore the possibility of employing an X-Y addressable flat plate display is a definite consideration. In addition an audio alarm coupled into the aircraft intercom system is included.

3. CONCLUSIONS

Based upon the concept feasibility programs [Fairchild Camera Corp., "76] recently completed we have experimentally verified the SSA concept. This result combined with a simplified algorithm has demonstrated a wire recognition capability even in a very cluttered background. These results coupled with the one dimensional description of a wire offer a potential solution to the wire detection/recognition/avoidance problem under low light-level conditions. This system, the WOWS, utilizes modern advanced technology and can reasonably be projected to provide this wire avoidance function in a small, lightweight, and low cost system. The WOWS has the further advantage of being totally independent of human variables associated with object recognition. This will

result in a system with rigorously predictable characteristics and a system which does not increase the work load of the aircraft crew when operating NOE.

ACKNOWLEDGEMENTS

REFERENCES

Amelio, G. F. "Charge Coupled Devices" Scientific American 230 22 Feb 1974.

Barbe, D. F. Proc IEEE 63 38 Jan 1975.

Barbe, D. F. IEEE Trans ED ED 23 177 Feb 1976.

Boyle, W. S., Smith, G. E. "Charge Coupled Semiconductor Devices" Bell System Tech J. 49 587 Apr 1970.

Fairchild Camera Corp. Syosset, NY. "CCD Format and Scanner" Contract DAAB07-76-C-0881 (1976).

Fairchild Camera Corp. Syosset, NY "Single Site Activation Logic and Display" Contract DAAB07-76-C-0927 (1976).

Jespers, P. G.; Vanderwiele, F.; White, M. H. Solid State Imaging NATO Advanced Studies Institute Noordhoff, Leyden 1976.

Kleider, A. "An Experimental Evaluation of Gated Low Light Level TV For Wire Obstacle Detection" ECOM Tech Report ECOM 4-321 May 1975.

Kleider, A. Proc 9th Army Science Conf. Pg 149, West Point, NY June 1974.

Sequin; Tompsett Charge Coupled Devices Academic Press NY, 1975.

TABLE I

Wire Detection Techniques

SYSTEMS	MODE	RESULTS AND COMMENTS
RADAR 60 MHz → 95 GHz	ACTIVE	• DETECTION GOOD ONLY WHEN NORMAL TO WIRE • SPATIAL LOCATION OF WIRE VERY POOR FOR REASONABLE SIZE ANTENNA • GROUND CLUTTER & FALSE ALARM RATES VERY HIGH • CONCLUSION --- NOT USABLE
ELECTROMAGNETIC FIELD DEVICE	PASSIVE	• DETECTION & LOCATION POOR • CONCLUSION --- NOT USABLE
LASER LOTAWS (10.6μ ; CO_2)	ACTIVE	• DETECTION & LOCATION GOOD (3cm, DIA WIRES AT 1.5Km) • DETECTION AT 60 DEGREES OFF NORMAL • MULTIFUNCTION CAPABILITY • FLIGHT TESTS OF EXPLORATORY DEVELOPMENT MODEL • CONCLUSION --- A VIABLE TECHNIQUE
LIOD (1.6μ Nd-YAG)		• LIMITED RESULTS. LOW REP RATE
GATED LOW LIGHT LEVEL TV (GaAs Illuminator)		• DETECTION & LOCATION GOOD (3mm DIA WIRE AT 700 METERS) • DETECTION AT ANGLES 60° OFF NORMAL • HIGH COST • CONCLUSION --- A VIABLE TECHNIQUE BUT LIMITED BY HUMAN FACTORS
WIRE OBSTACLE WARNING SYSTEM (GaAs Illuminator)	ACTIVE	• IN EXPLORATORY DEVELOPMENT

PASSIVE ──────── | ELECTRIC FIELD SENSOR

RADAR &
EM SYSTEMS

ACTIVE ──────── 60MHz RADAR
9 GHz RADAR
95 GHz RADAR

INCOMPATABLE

PASSIVE ──────── FLIR
LOW LIGHT LEVEL TV

OPTICAL

ACTIVE ──────── GATED LOW LIGHT LEVEL TV (GaAs Illuminator)

LASER ──────── LIOD (1.06μ Nd YAG)
LOTAWS (10.6μ CO_2)
WOWS (GaAs Illuminator)

FUNCTIONAL
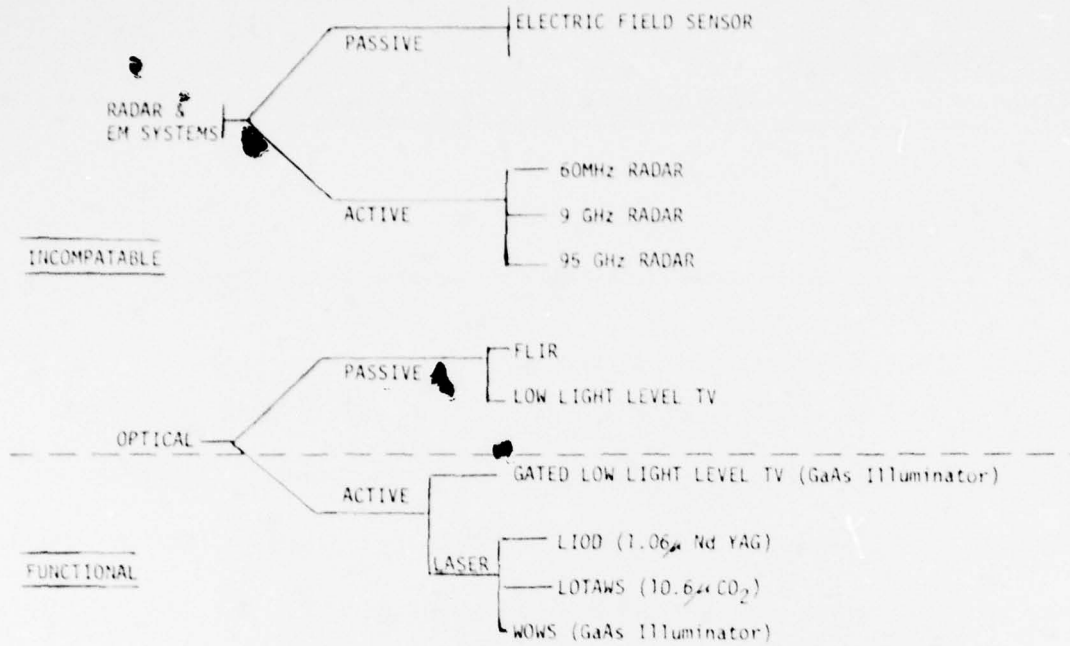
Fig.1 Wire obstacle detection techniques tree

Fig.2 GL^3TV feasibility test
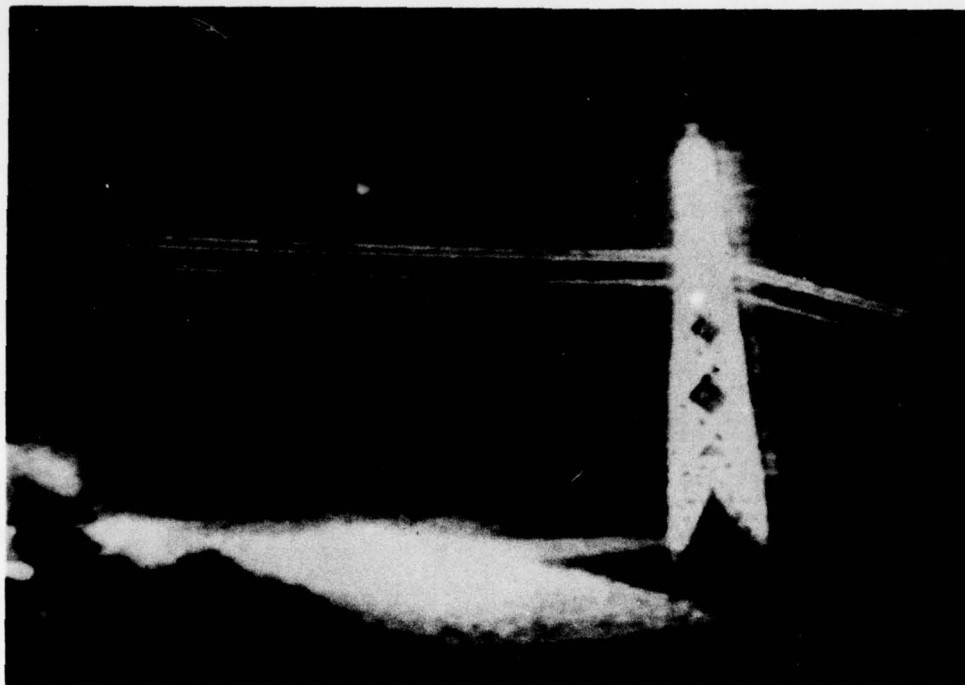
Power Transmission Lines (Active Mode)
Range: 2250 meters   Wire: upper; 1/4" Al clad
Camera FOV: 5.6°   lower; 1" Al clad
Incidence Angle: 0°   Laser Power: 0.6 watts
Eff Contrast: 71% (1" wire), 38% (1/4" wire)

Fig.3   GL³ TV wire detection
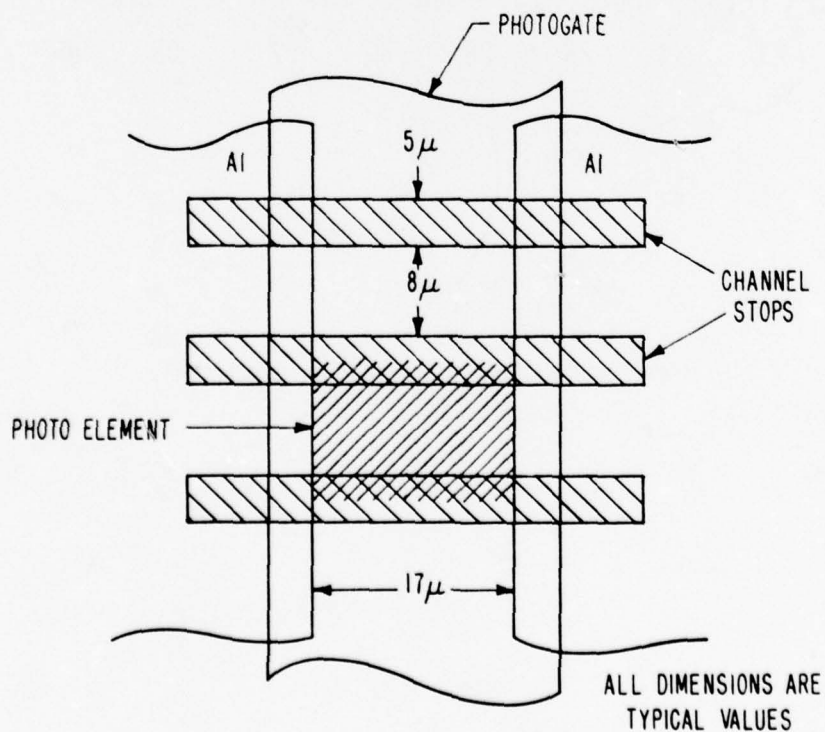
Fig.4   CCD photoelement dimensions
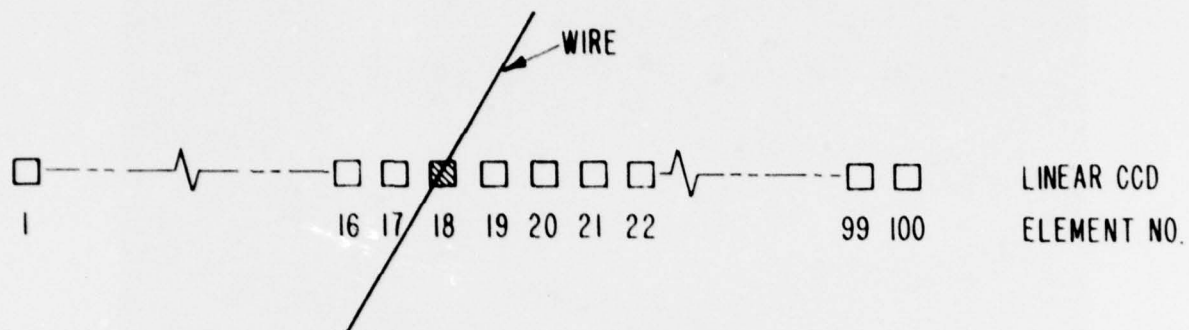
Fig.5 Wire in image plane of a linear CCD array
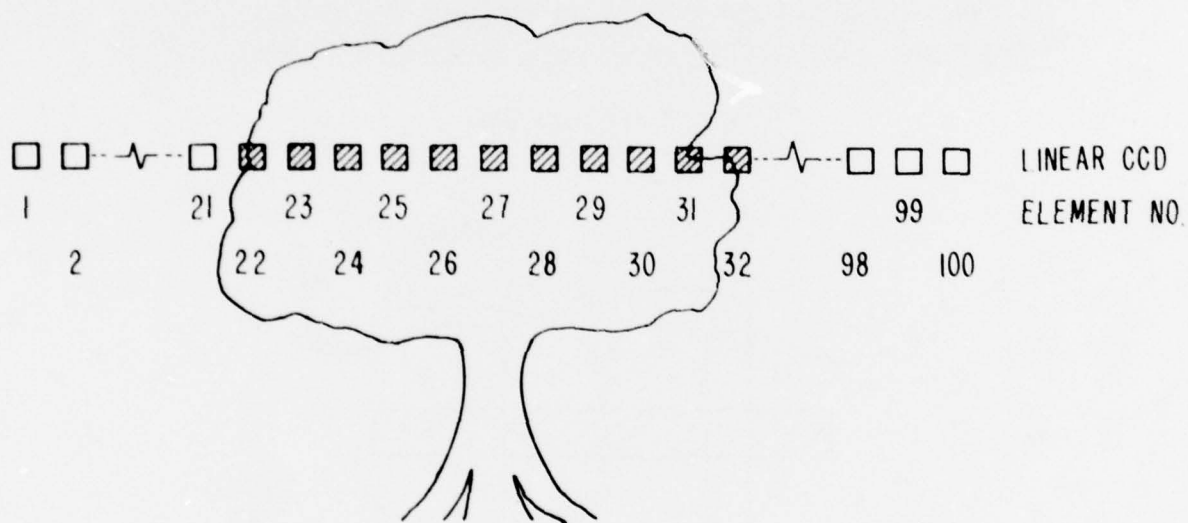
Fig.6 Non-wire obstacle in focal plane of linear CCD

A THRESHOLD

B RAW VIDEO
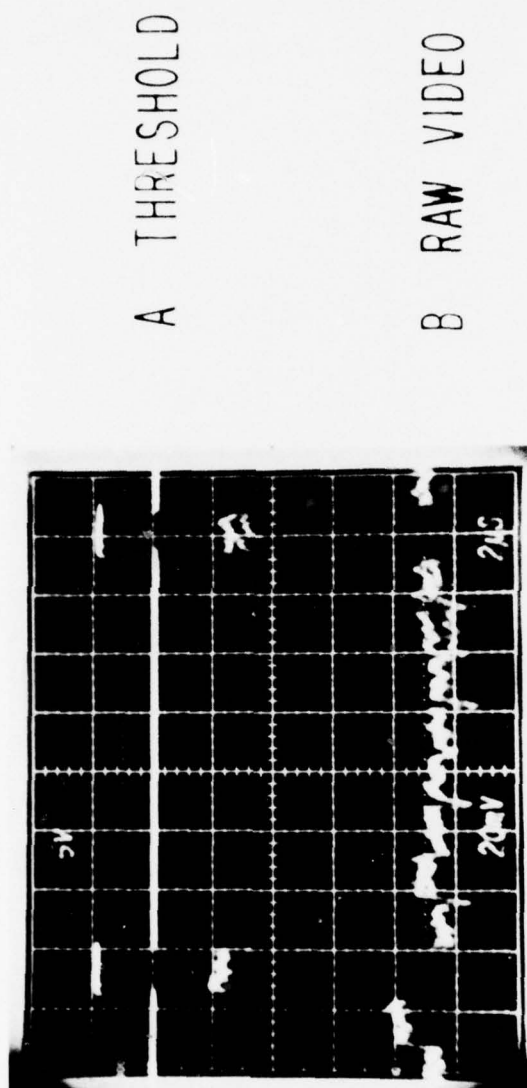
Fig 7 Experimental single site activation

1.75 x 10⁻³ RADIANS

0.175 RADIANS

10⁻⁴ RADIANS

0.524 RADIANS
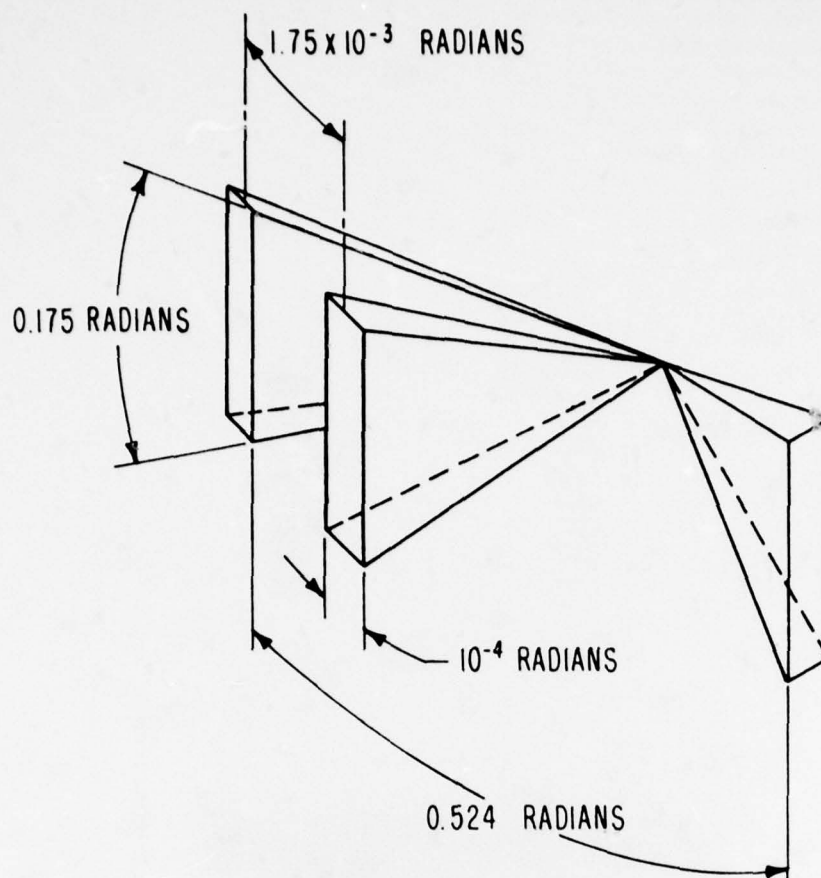
Fig.8   WOWS fields of view

M SCANS

N ELEMENTS
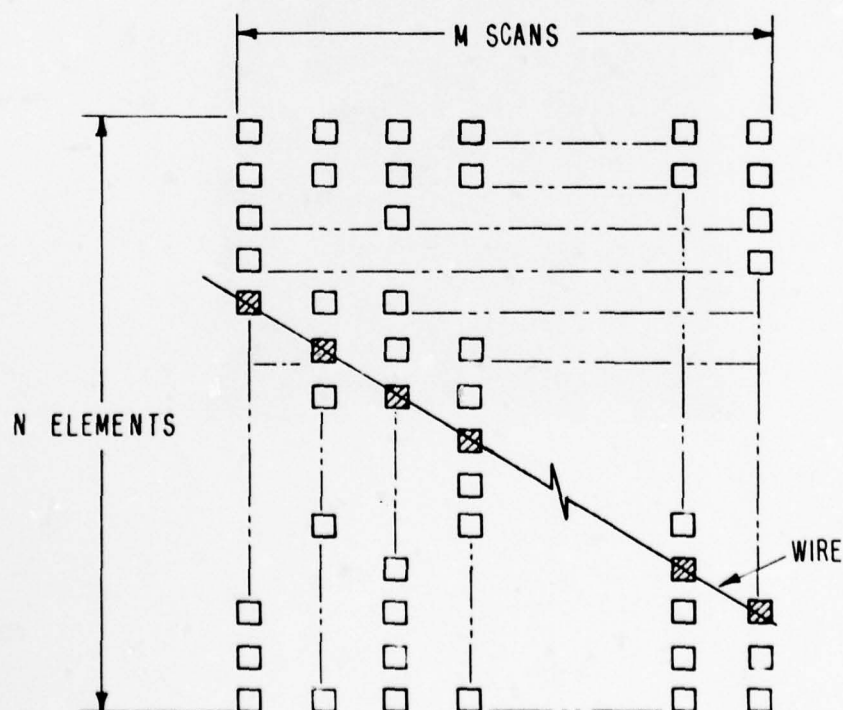
WIRE

Fig.9   WOWS detection matrix

Fig.10(a) Wire display 1% background

Fig.10(b) Wire display 2% background

Fig.10(c)   Wire display 5% background
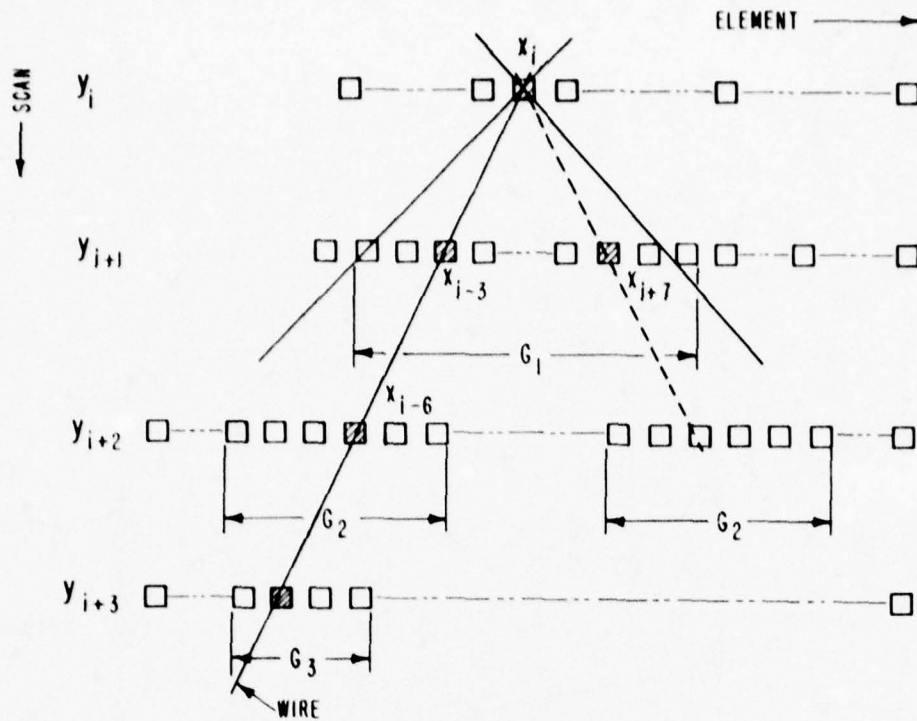
Fig.11 Predictive gating

t_s - FRAME TIME

θ_d - ANGULAR COVERAGE PER DECISION
ELEMENT - FOV (I) · n SCANS

θ_t - TOTAL FOV

$2\Delta R = C\tau$ where τ - ILLUMINATOR PULSE
WIDTH
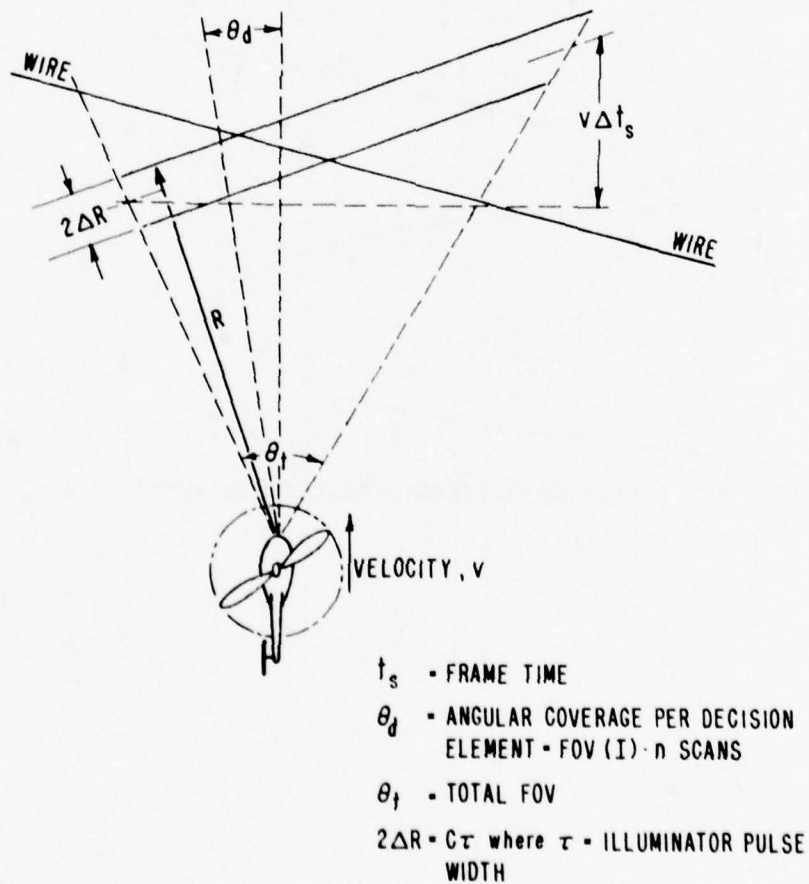
Fig.12 WOWS "longitudinal" scan
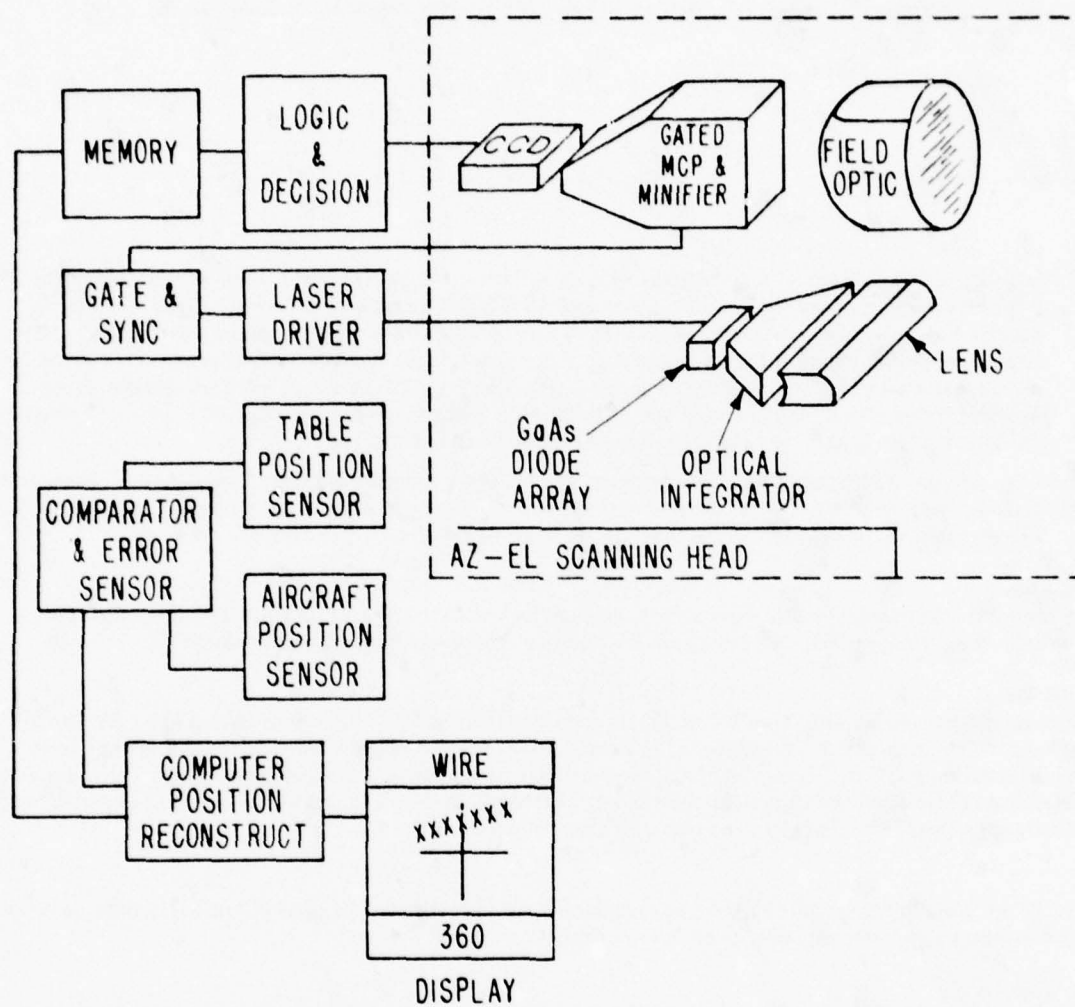
Fig.13   Wire obstacle warning system schematic

DISCUSSION

Bernhard

1 — How do you handle the case where a wire is positioned between the pixels; or is it rejected?

2 — In the case of 8-point or 15-point wire algorithms, are any missed points allowed?

3 — What impact does the above have on detection probability?

Author's Reply

1 — A wire will cover more than one pixel when it is spatially oriented at an angle less acute than the corresponding aspect angle of the individual CCD element or when it is laterally displaced, as shown below.
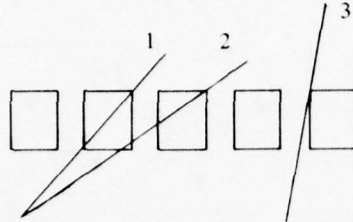
In wire 2 above, on each scan more than one pixel will be activated. Therefore, if we wish to cover such occurrences, the single site activation algorithm (010) would have to be modified to a (0110) in order to provide a dual-element equivalent. This would in turn increase the number of objects which could possibly fit the recognition pattern. The impact of such a change would require a detailed examination of the effects upon the desired system purpose and operation. Our contention has been that there are other methods available for detecting and locating larger obstacles while there are no reliable methods for wires. Therefore our intent is to see wires, perhaps at the expense of missing larger objects.

2 and 3 — The geometric algorithm which reconstructs the linear properties of a wire from the sets of single site activations found in subsequent scans permits an adjustable number of drop outs and misplaced detections. These errors are continued in establishing the decision criteria.

J.J.Stapleton

In Europe another threat besides wires is the round aperture of TV trackers. Can your algorithm be adapted from the line-spread function to a point-spread with use of your image motion compensation?

Author's Reply

Without getting into sensitive areas let me simply comment that the WOWS is strongly based upon the Single Site Activation concept. It is possible that an algorithm could be developed to provide recognition of a line-spread function geometry. However, at this time the WOWS does not appear to offer more than a germ for a solution to the problem which you have addressed. Further discussions and comments should be undertaken at a more appropriate time and place and would be secret in nature.

Robin

Quelle est la sensibilité de votre système au brouillard et aux aérosols. Quelle est la distance maximale que l'on peut espérer de ce système compte tenu du brouillard et autres phénomènes?

Author's Reply

The WOWS is an exploratory development program designed to test the feasibility of the concepts involved in the detection and recognition of wire obstacles. This system has been designed to operate in clear weather and under lighting conditions appropriate to starlight to dawn-sky conditions. Based upon atmospheric transmissivity data cited in ECOM Technical Report 7023 (entitled Atmospheric Optical Environment, by M.L.Vatsia, dated September 1972) in a moderate haze where the visibility was 1 km the transmissivity would decrease by 3db. Based upon present design considerations this would decrease our range capability to approximately 200 meters.

However, under limiting visibility conditions the helicopter velocity would most likely be similarly reduced. This in turn would reduce the required detection range. Therefore, while the system has not been conceived as an all-weather device it would appear to be useful in weather conditions which were not ideal.

In addition, the development of IRCCD's is being closely followed. If and when 10.6μ devices with an appropriate physical size are made available the WOWS can be combined with a CO_2 laser-pulsed illuminator to provide improved all-weather capabilities. However, it is required that the IRCCD provide the equivalent one-dimensional high resolution.

Carlier

Quel est le coût de cet équipement?

Author's Reply

About 25,000 dollars a copy.

A CCD DIGITAL IMAGE STORE

D. M. Balston and P. R. Samways
Plessey Radar Research Centre,
Southleigh Park House,
Havant, Hants.
U.K.

SUMMARY

This paper describes a digital image store, based on CCD technology, which has been specially designed for an image processing system. The latter, termed IDP 3000, has been introduced to fulfil a requirement in the field of Earth Resource Surveying and the design of the store reflects the requirements of this specific application area. The store has however proved extremely flexible and totally unrelated projects are currently finding it of use.

The paper first introduces the requirements and then describes the specific solution adopted. Particular attention is paid to the relationship with a standard 625 line TV signal. Each function available in the store is discussed and the method of implementation is described. The final section gives an indication of the reliabilty of the store and discusses probable future extensions, closing with a brief description of its application to simulator development.

1.   INTRODUCTION

The expanding range of imaging sensors and the concurrent improvement in platform technology has created a considerable increase in the work load of earth science photo-interpreters. Already over-stretched these experts are finding that they cannot keep up with the ability of remote sensing technology to provide them with images. Several systems have been developed to aid the interpreters in these tasks, and this paper describes the development of the image store module associated with one such equipment, the IDP 3000.

The main function of the IDP 3000 is to provide an interactive image processing capability which allows the interpreter to define areas of known ground cover. The system then extrapolates over the whole image area and provides statistical and cartographic information for incorporation in a database. To fulfil these functions the interpreter must be able in real time to undertake contrast enhancement, colour enhancement, area definition, theme classification and communication with a host computer.

One of the central features of the system is a specially developed image refresh store based on CCD technology. The ideal store for this application would be a fully random access memory operating at 100 nS. Such a solution was considered prohibitively expensive however, particularly as the number of images needed to be stored at any instant was expected to be high. An alternative approach was there-fore sought which was less costly and yet retained the speed and flexibility necessary for image processing. The cost per bit of CCD memory was predicted to fall very rapidly compared with other fast solid state stores and this, coupled with their suitability for TV related applications led to their choice for this application.

This paper describes the store in terms of the requirements, how they have been met, the current capability and possible future extensions.

2.   STORE REQUIREMENTS

2.1   The Image Processing Environment

The image processing environment requires that several different types of processing are available to the earth scientist interpreter. For example a general purpose computer is generally needed to perform the more complex tasks and allow the development of new processing algorithms. Real-time processing is essential also because a considerable number of the interpretation tasks can still only be undertaken manually (or more accurately, by eye) and optimisation of the displayed data is essential. This cannot be achieved if the delay between the setting up of a process and its subsequent display is too long. In fact considerable effort has been devoted to the ergonomics of these optimisation processes and responses of 0.1 seconds are deemed desirable (BALSTON, D.M., 1976). Only a limited range of these processes can be supplied at interactive speeds without severe cost penalties. It is not feasible however to rely on a general purpose computer if processing times are to be commensurate with the timescales associated with many operational tasks. An agricultural inventory of potential crop yield would be useless if it did not produce its results until after the harvest. Thus a third type of processing, not interactive yet special purpose and much faster than a general purpose computer, is necessary.

This latter processing has a particularly important role to play in the continued development of the IDP 3000 because it is the means by which an additional important processing capability will be introduced in the near future. All systems of this type are currently capable only of processing each picture element in isolation and thus cannot make use of such features as texture, shape, closure, continuity, repetition etc. These characteristics are known as spatial features and the introduction of a capability to handle these is essential to the success of image processing as a discipline.

The store has to interface to all these processing regimes. Figure 1 presents the basic structure of the IDP 3000 and shows how the store itself is a central element. The Interaction Control Processor interprets the user's commands and routes them to the required processing or store modules. It is also capable of writing to the store thus enabling image statistics to be displayed graphically and alpha-numerically on the colour TV monitor. Photographic source images are input via the monochrome TV camera while images on computer compatible tape (CCT) are input via the host computer. Image output can be photographic, directly from the TV screen or by means of an off-line photowrite facility via CCT. In order to maintain processing consistency and repeatability all functions, excluding the camera and monitor, are implemented in digital electronics. The video rate processing provides such functions as image ratioing, algebraic channel combination, contrast enhancement and colour classification.

2.2   The Requirements

The requirements on a store to support the system structure described above are many and varied. This section describes the important ones, other facilities which have been incorporated in the design are discussed in Section 3.

The store must obviously be capable of receiving data directly from a TV camera and preferably storing a single frame rather than building up an image line by line over a period of time. The data supplied to the video rate processing must comprise a standard TV signal in 625 line format. This allows commercially available studio quality TV equipment to be used wherever possible. Data must also be accepted from the host computer, special processors, other stored images and the interaction control processor. Data must be sent to all but the last.

The spatial resolution should be at least commensurate with the best colour TV monitors. This leads to a 512 x 512 format.

The store should be modular in structure, so that many images, or several colour bands of one image, can be accommodated and extra channels can be added with ease. The intensity resolution of each channel should also be modular so that the capacity can be matched to the data source.

One of the important facilities in the image processor concerns the ability to classify the data into user defined themes. For example crop types for an agronomist. This facility can produce several single bit images which have subsequently to be combined to produce multiclass images, thus it must be possible to write to individual intensity bits in any channel without affecting the other bits.

It is desirable that the interpreter be able to compare processes side by side on the display screen, thus enabling him to optimise the display for specific tasks. This can best be achieved if the data in each half of the screen is identical and thus a further requirement is the ability to replicate part of the image in another area of the display. This facility should be available in both the horizontal and vertical dimensions.

The special processors discussed in Section 2.1 need particular mention because the majority of them will require that the data be provided in the form of a small two-dimensional window within the image. They will then scan this window across the complete image, determining spatial features at every point and writing the answers to other store channels. This spatial processing data bus is essential to the success of the system and must operate at high speed.

3.   STRUCTURE OF THE STORE

3.1   General Description

A block schematic of the store is shown in Figure 2. A central controller provides master timing and synchronising data to allow the generation of standard 625 line interlaced TV signals by each channel when required. It also provides the interfaces with the controlling system, the data sources and the special processors.

The store channels each have independent local timing and control circuits. This enables them to operate independently of their neighbours and also reduces the number of fast control signals which have to be bussed around the system. Each channel can at present contain up to 8 intensity bits although a very simple modification would extend this to 20. Each bit is contained on a single board thus giving considerable modularity and flexibility of use. Each board, in addition to supplying the standard video output, can supply a "regular video" signal. This signal is used for functions involving transfers to non-video rate processors and is described in more detail in Section 3.3.

The system currently in use is designed for a total capacity of 8 channels. Extra channels could be accommodated by a slight increase in the output switching complexity.

3.2   Functional Description

Each channel can operate in one of eight modes. Although the local controls are independent, there are operational constraints imposed by the central controller which limit the range of modes which can be in operation at any instant. These limitations are shown in Table 1. Selecting a specific mode for a channel limits other channels to those modes defined by ticks in the selected mode row.

The specific modes are discussed in more detail below.

Mode 1:   Refresh
The channel supplies standard interlaced TV signals to the processing and display unit via the switched output bus. This bus contains 8 video lines and each channel can be switched to any lines as determined by an output configuration register set in the channel controller. A centralised bus conflict detector prevents any attempt to enable data from two sources onto the same line.

Refresh is the basic operating mode of the system and all other modes return automatically to refresh when complete.

In refresh mode the displayed signal can be further manipulated by the channel controller as follows,

   (i) Horizontal replication - half the horizontal extent of the displayed signal can be shifted to the left hand edge and repeated from the centre.
  (ii) Vertical replication - as for horizontal replication but in the vertical direction.
 (iii) Reduce resolution - either of the interlaced TV fields can be repeated to give a half resolution display in the vertical direction, thus doubling the apparent capacity of the store. The horizontal equivalent of this would require a slight extension to the control functions.

Mode 2:   Video load
A complete frame of video data, from an external source or another channel, is loaded into the defined channel(s). The delay of the incoming data with respect to the data already stored is accommodated by the channel controller to maintain image registration.

Facilities available in the video load mode comprise,

   (i) Bit write mask - enables individual bits of the intensity range to be accessed.
  (ii) Erase - writes zero in every location.
 (iii) Reduce resolution - enables two independent images, each of half the standard vertical resolution, to be loaded into the separate fields of the store.
  (iv) Mirror - reverses the x scan direction thus providing a mirror image prior to storage.

Mode 3:   Load computer
This mode allows data stored in any channel, or group of channels, to be transferred to a host computer. In the system developed the host is a Prime 300 mini computer but transfer is via a normal 16 line DMA link and interfacing to other similar hosts would be straightforward.

Data is normally loaded into the interface during field blanking periods and there are thus no visible effects if the same channel is also being used for refresh. Loading can however be performed during display time if the extra speed is required.

Multiple channels can be transferred in parallel on a line by basis. Only those lines required need be transferred.

Mode 4:   Computer load

This mode is practically the inverse of mode 3 with similar constraints and facilities. When multiple channels are being transferred however a hardware circuit in the controller minimises the software sorting and disc access times which could otherwise cause considerable overheads on the transfer times.

It is not necessary to transfer the whole 512 x 512 image. Transfers are undertaken on a line by line basis and the horizontal start and finish coordinates can be defined to 4 bit accuracy thus allowing transfers in multiples of 32 picture elements.

Mode 5:   Window source
A 16 line RAM buffer allows a 16 x 16 data window to be scanned across the image under external processor control. When scanning is complete an external command can cause the buffer to be reloaded. It will generally be reloaded with the group of 16 lines one line removed from the first, thus allowing the window to scan down the image.

Mode 6:   Asynchronous line load

This mode enables data to be written into the store from an external source which is asynchronously timed with respect to the store timing (e.g. a thermal line scanner). The incoming data is written on a line by line basis. The process is very similar to Computer loading, Mode 4,

with the important distinction that the controller cannot in general tell the asynchronous source when data can be accepted. This naturally imposes limitations on the data rate of an asynchronous source and in general this is 4.7 mS per new line of data.

In conjunction with this mode is the facility for accepting a continuous string of lines (subject to the rate limit) and continually inserting these lines at the bottom of the screen. The displayed image then rolls up the screen at a rate determined by the incoming data. This roll can be stopped at any time to freeze a frame for detailed examination.

Mode 7: Cursor load
All modes discussed thus far accept whole or partial lines of data in blocks no smaller than 32 picture elements. This is obviously inconvenient if only small changes are required to the stored image, for example the insertion of annotation or lines derived from a coordinate entry device. The cursor load mode allows changes to single picture elements to be made by the interaction control processor.

Mode 8: Window receive
This mode is used in conjunction with the window source mode to accept the processed data. All the facilities associated with computer load are available.

3.3   The 1-bit Store Board

The CCD chip used is the Intel P2416 (INTEL, 1977). This comprises 64 sets of digital shift registers each 256 bits long. These dynamic devices must be shifted at a frequency in the range 0.1 MHz to 1.3 MHz. Shifting signals are provided by Intel 5244 drivers from a four-phase clock supplied by the central controller.

The store board is designed to hold one intensity bit of a 512 x 512 digitised TV frame. A normal broadcast TV system displays approximately 576 of the 625 lines, the remaining lines being used for frame flyback, colour burst etc. The 64 additional blank lines are generated by the system but removed from the display by extending the height of the raster on the monitor.

In order to maximise the storage efficiency and reduce the operating frequency of the CCD's the standard TV signal is buffered on entry to the board and the line blanking period ignored. Figure 3 shows that data is read into the 512 element buffer at 10 MHz rate (i.e. 100 nS per picture element) over the 51.2 μS of the visible line period. During the 64 μS total period of the subsequent line this data is read into the CCD's at an 8 MHz rate. Double buffering allows data to be stored as it arrives from the TV camera and thus a full frame of video data can be stored in 40 mS. Readout to provide refresh signals for the processing and display are obtained by inverting the process.

The 8 MHz "regular video" signal thus obtained still has a data rate considerably higher than a single device can accept. On each board therefore 16 devices are multiplexed together, each device operating at an overall data rate of 0.5 M bits/S.

Figure 4 shows the structure of the CCD data block in more detail and shows the organisation of TV lines within the CCD elements. The 16 CCD chips are stacked side by side along the X axis. Shifting is shown along the Z axis, recirculation of the data occurring automatically. At a given shift position in the refresh mode the 16 values are read into the output serialiser and hence to the buffer. At the same shift position the 16 values from the next register are read and serialised. The store is then shifted and the process repeated. Thus 32 picture elements are obtained for every shift. Shifts occur at 4 μS intervals and one line of data occupies 16 shift positions. There are 256 shift positions and thus a complete recirculation of the store takes 16 lines, or 1.024 mS. The set of two registers providing these 16 lines of data is termed a shelf.

Interlace and field blanking cause an additional complication to the control functions. Each field comprises 312$\frac{1}{2}$ lines of data. In the store this is taken as 313 in field 1 and 312 in field 2. There are only 256 lines of data to be stored however and thus the system must wait 57 lines between fields 1 and 2 and 56 lines between fields 2 and 1.

In order to simplify overall addressing and to ease the provision of such functions as computer load and window source it is desirable that the two fields be stored in phase, i.e. the first picture elements of each field should occupy the same shift position. This requires that the store must recirculate an integral number of times between data fields. This is achieved by supplying a period of double speed shifts during field flyback. Between fields 1 and 2 this lasts for 7 lines (448 μS), thus the store is shifted through 50 lines at normal rate and 14 lines during the double speed period to give a total of 64 lines or four complete recirculations. Between frames 2 and 1 the double speed shifting lasts for 8 lines (512 μS)

3.4   The Central Controller

Although the control of the operating mode resides within each channel the central controller is kept informed of all currently operating modes. This is essential for the timing signals for every mode are generated centrally and distributed to every channel. Local mode control registers are used to select the timing signals required for the selected mode.

The centrally generated signals comprise the normal and double speed shift signals, the line replication command and the source data selection commands. The shift signals are rather more complicated than outlined in the preceding section. The double buffer on the input to the store board introduces a two line phase difference in the CCD recirculation between the refresh and video load modes. In order to maintain synchronism between the two video signals the phase of the channel being loaded must be advanced by two lines prior to the load taking place. Signals enabling the generation of single and double speed shifts are stored in PROM for each mode requiring them. Standard TV synchronising signals comprising syncs, blanking and field equalising pulses are also enabled from PROM. A current shelf address increment command is distributed after every 16 lines to each waveband and used by those in refresh and video load mode. The remaining signals generated centrally concern the interfaces with external data sources and comprise mainly addresses and clocks.

3.5   The Channel Controllers

Each channel controller contains 7 registers which define and control the alternative modes of operation.

The output configuration register allows the channel to be switched to any one or more output lines, up to a maximum of 8. A requirement of the image processor which uses this store is the ability to switch output paths during the visible frame time and thus this 8 bit register is repeated 16 times to give an extremely flexible real-time output control capability.

The second register defines the mode of operation and the number of other channels in the same mode. This latter information is essential when multichannel transfers are taking place between store and host computer. Register 3 defines the order in which multi-channel data is loaded.

The remaining registers define the internal operation of the channel. Reduced resolution, erase, mirror and roll are controlled by register 4 while 5 allows the individual bits in a channel to be masked out prior to loading. Registers 6 and 7 define the start addresses for horizontal and vertical segmentation.

4.   PERFORMANCE AND FUTURE ENHANCEMENTS

4.1   Reliability

Three 8 bit channels have been in service now for 10 months and have clocked up over 1600 hours. A further channel, also 8 bits, has been working for half that time. Each channel contains 128 CCD chips and thus the total number of device hours to date is over 700,000. In that time only one device has failed. This figure, although not statistically significant does nothing to undermine the manufacturer's claim of 2×10^6 hours MTBF at 55^oC.(INTEL, 1976).

The system itself contains approximately 2000 TTL devices and, apart from one catastrophe when an interface failed during development, only 4 devices have failed in service. The $2\frac{1}{2}$ million device hours attained to date is less than would be expected for 4 failures but the equipment has been operating under severe experimental testing conditions during this period. It is not yet possible therefore to draw any valid conclusions concerning system reliability.

4.2   Flexibility

The system of which the store is a central feature has proved extremely powerful and one of the main reasons for this has been the range of facilities which have been built into the store. Although designed as an earth resources analysis system it has already been reprogrammed to perform tasks totally unrelated to those for which it was designed. One of these is worthy of note as it uses the capability of the store to modify its characteristics in real-time. By storing a monochrome image of a typical battlefield scene in one channel, constructing a foreground section by means of the earth resources scene classification facility and storing this in a second channel, a potential simulator for weapon sights can be produced. A third channel can store real projections of the target and these can be selected automatically by the system controller as the target traverses the scene. The video processing developed for earth resources allows the data in any channel to control the selection of data onto the display and this, in conjunction with the ability to move one channel (i.e. the target) with respect to the others produces a potentially powerful and realistic simulator. This single example has made use of the following facilities,

    (i) Video load - to obtain the original images of scene and target.
   (ii) Video refresh - to display scene and target
  (iii) Load computer - to allow editing of the images and the foreground mask.
   (iv) Computer load - to restore edited images prior to use.
    (v) Roll - to enable smooth vertical movement of the target
   (vi) Bit write mask - to associate target switching data with target image data.
  (vii) Horizontal replication - to produce horizontal movement
 (viii) Output configuration - to select target, background or foreground.

4.3   The Future

Two areas in which the store could usefully be extended have already arisen. These concern its use in the earth resources role, and refer in the first instance to the need for a closer look at a portion of image without necessarily displaying additional data. This image expansion capability allows the earth

scientist to determine more precisely the exact position of a feature of interest. The second related need is for the storage of larger scenes with higher resolution. A typical Landsat scene for example contains approximately 2,500 lines each of 3200 picture elements.

The 16 line recirculation characteristic makes functions such as image expansion difficult to achieve. For example, if it was desired to expand the top left quadrant of the display so that it filled the screen, it would be necessary to display line 1 twice then line 2 twice etc. When line 2 was needed however the store would have circulated past it and almost a whole recirculation would be needed before it was available. Thus image expansion cannot take place in real time direct from the store.

It is proposed to overcome this problem by devoting an extra channel to the expanded image and writing to this from those store channels to be expanded. Expansion would be limited to even integral factors up to 16. Because an expansion of 2 means that only 25% of the original data is needed for the display, up to 4 channels of expanded image can be stored in a single additional channel. Many of the control functions would be unnecessary in this expansion channel and the extra demultiplexing needed to allow the four expanded images to be simultaneously available would be comparatively simple.

The cost per bit of CCD memory is likely to decrease substantially only with the introduction of larger memory capacity on a single chip. 65K chips have already been announced. This will probably be accommodated in the current store architecture in such a way as to increase the image resolution. This will allow a display to pan over the stored data providing a video window on a scene two or more times larger in each dimension. It remains to be seen whether this approach would also allow an overview (i.e. the reverse of expansion) to be provided or whether the expansion channel would serve this function also.

5.   CONCLUSIONS

This paper has described the development of a digital image store using CCD technology. The environment and system requirements have been presented and the structure and performance of the store described in detail.

The store has proved extremely successful and the facilities which have been incorporated make it a very flexible system. It has already been used for other tasks.

The cost however is not significantly less than equivalent RAM refresh stores and although this is due in part to the increased control requirements the main reasons are associated with the continued fall in price of RAM and the larger quantity production of systems. It is anticipated that the introduction of higher capacity CCD memories will allow the store to become extremely competitive.

ACKNOWLEDGEMENTS

Selected Mode		ALLOWED MODES FOR OTHER CHANNELS							
		1	2	3	4	5	6	7	8
1	Refresh	√	√	√	√	√	√	√	√
2	Video Load	√	√						
3	Load Computer	√		√					
4	Computer Load	√			√				
5	Window Source	√							√
6	Asynchronous Load	√					√		
7	Cursor Load	√							
8	Window Receive	√				√			√

TABLE 1: Limitations on channel mode selection

REFERENCES

BALSTON, D.M. 1977 - General systems considerations and the IDP 3000
Plessey Radar Research Centre Report
No. 17/77/PO33U, February, 1977.

INTEL 1976   Reliability Report RR-11
Intel Corp. 3065 Bower's Ave., Santa Clara
CA 95051, May, 1976.

INTEL 1977   General Product Catalogue
Ibid.
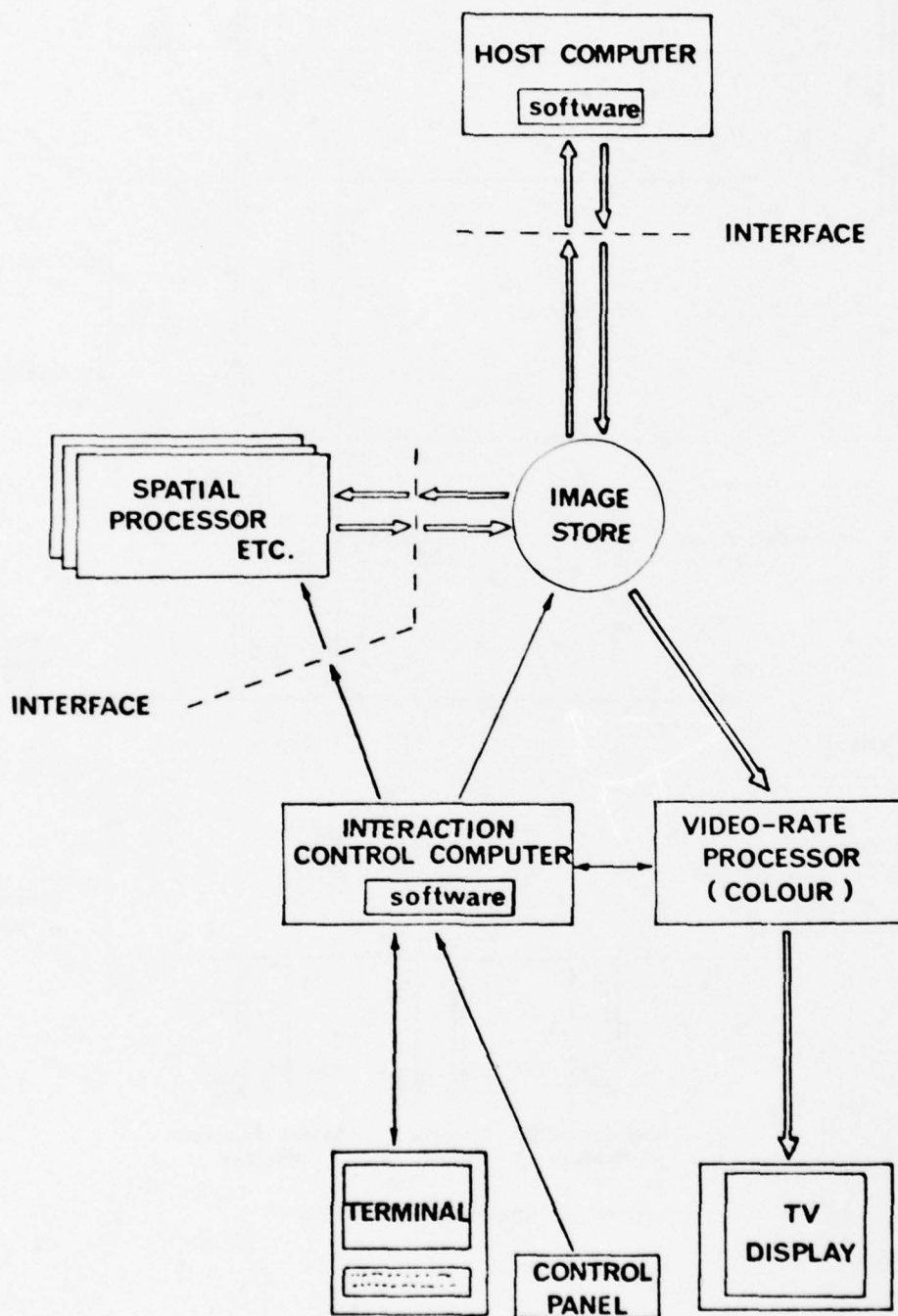
FIGURE 1 - IDP 3000 - BASIC STRUCTURE
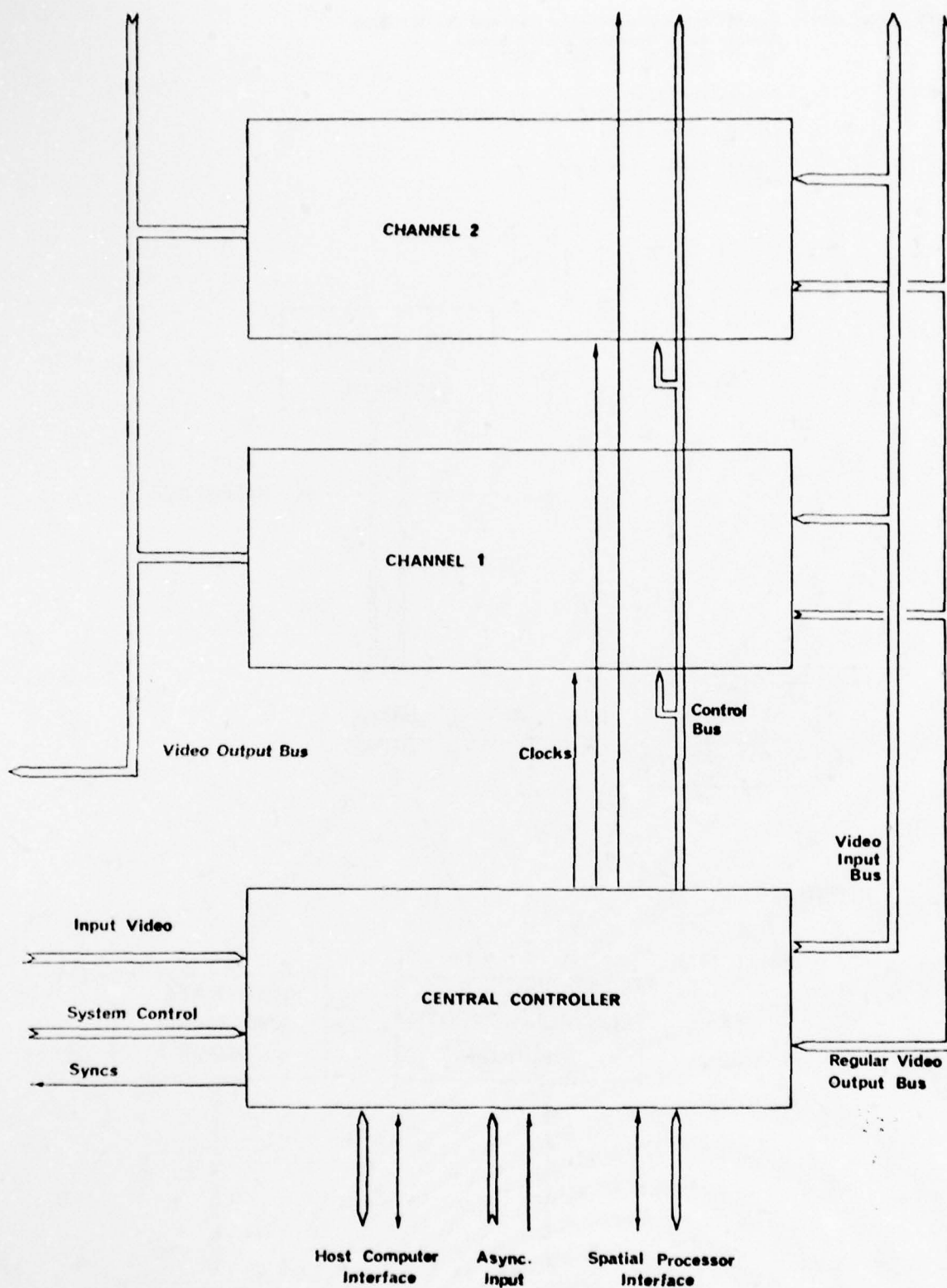
FIGURE 2 – BLOCK SCHEMATIC OF STORE
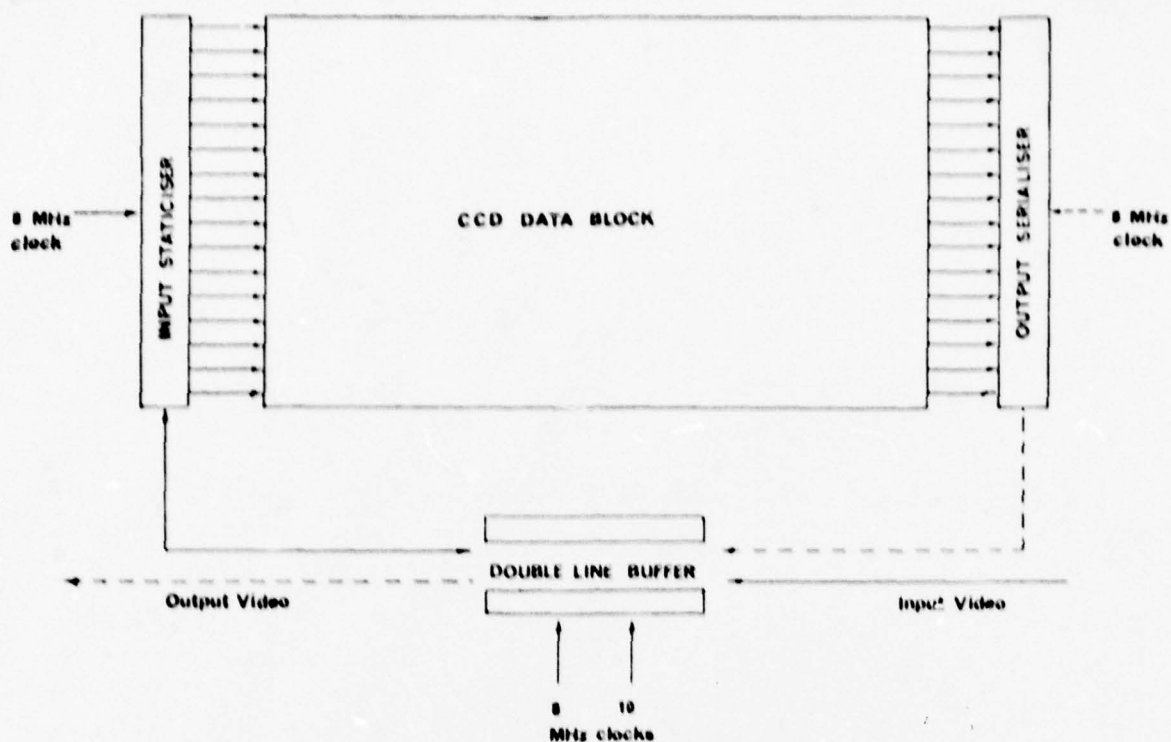
FIGURE 3 — SCHEMATIC OF STORE ACCESS CHARACTERISTICS
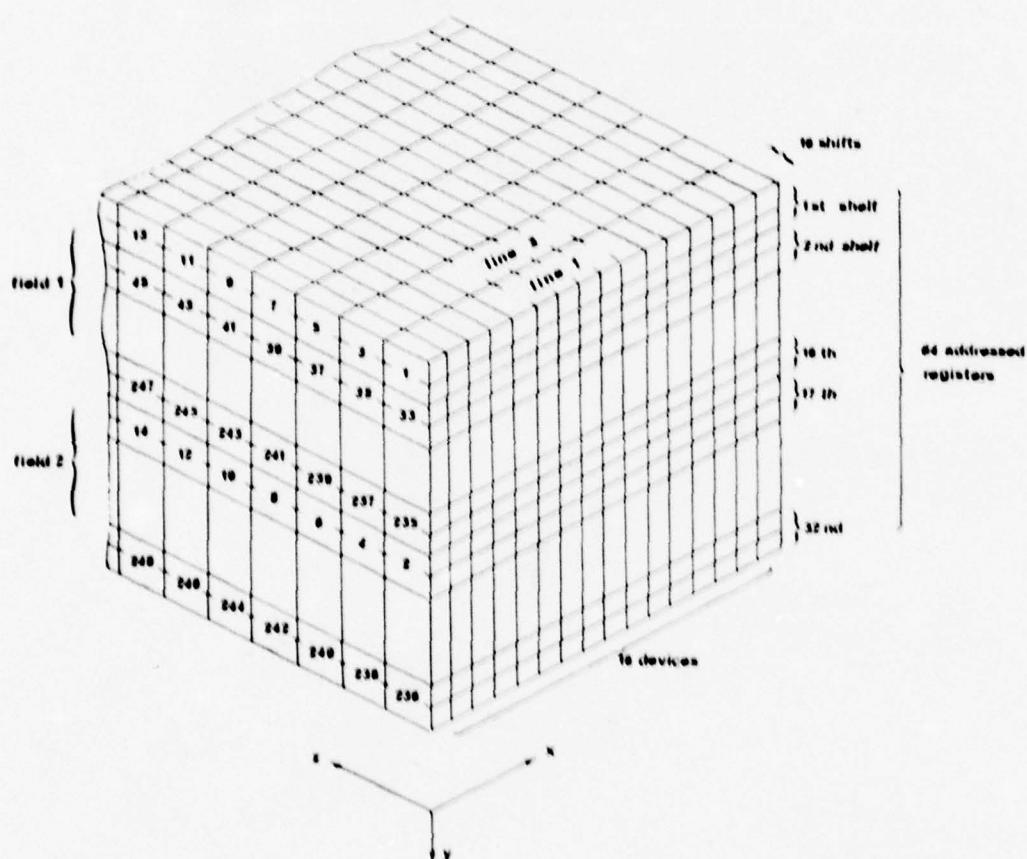
FIGURE 4 — STORE ORGANISATION

A CCD MEMORY CHIP FOR RADAR IMAGE PROCESSING

K. Rödde
E. Stein
H.-J. Wulf

AEG-Telefunken
Forschungsinstitut
Ulm, Germany

SUMMARY

Considerations for the design of a digital CCD memory will be discussed. Starting from commercial CCD devices, the design characteristics of a second generation CCD will be derived. The economically feasible capacity and its dependency on the technology lead to the definition of performance parameters and the chip organization. A SPS organization will be found most suitable and design considerations of SPS blocks will be discussed. Finally, the overall design of a 32 kbit CCD will be described.

1. Introduction

In surveillance radars /1/, /2/ there is great interest in semiconductor memories for two reasons. First, if a frame store is available at a reasonable price, it will be possible to use an image refresh memory instead of a storage display tube for scan converters resulting in less bulky and more reliable equipment. Second, a frame store could be used for MTI (Moving Target Indication) if received radar signals from two subsequent antenna revolutions are compared in order to detect tangentially flying targets.

In the first case a shift register memory is completely sufficient. In the second case, access to the memory is not in a completely sequential manner although there is strong emphasis on sequential access. Shift registers which are not too long will be sufficient, and no RAM is necessary. Thus, for both cases a CCD memory could be used if the limited flexibility is compensated by a lower cost as compared to RAMs. In the following two sections some reasons for the development of a second generation CCD chip are given. In section 4 we will discuss some aspects about cell area, memory block organization and power consumption. In section 5 we will compare different SPS block organization possibilities. Then in sections 6, 7, and 8 realized SPS blocks and the design of the 32 kbit chip will be explained and the performance discussed.

2. CCD memory systems and chips

Typical CCD memory systems for the above discussed applications will have capacities between 1 and 10 Mbit. Data rates range from 10 to 100 Mbit per second. A small memory system has been built with Intel 2416 CCDs and has worked very satisfactorily. The economic advantage over RAM memory systems however was not as large as would be necessary to make the CCD memory preferable. The following reasons could be identified: (1) The organization of the chip provides high flexibility which is not always required, but adds to overhead. (2) The storage capacity per chip is too small. (3) Clock driver requirements are high, because 4 high-capacitance clocks have to be driven. From these findings, the following requirements for second generation CCD chips can be formulated. These are: (1) Simpler chip organization, yielding higher packing density. (2) Higher chip capacity at essentially constant chip area to decrease cost. (3) Only two external clocks with a minimum capacitive load.

3. Design of a second generation CCD chip

The above statements have been at the starting point of digital CCD development. The first task was to define realistic performance parameters for a 32 kbit memory chip which would result in acceptable production yield and reliable operation.

Fundamentally the aim of every memory chip design is high capacity and low access time. The serial structure of CCD memories leads to a latency time, which passes before access to a specified location is possible. For a short latency time a serpentine organisation, which is shown in Fig. 1a, with short registers and high clock frequency is suitable. In this case overhead is high and consequently the capacity for a fixed chip size is low. The Serial-Parallel-Serial organization (SPS), shown in Fig. 1b, consists of a serial input register which is n_s bit long. The contents of the serial register is copied by one transfer pulse to n_s parallel registers, each n_p bit long. After n_p simultaneous transfers in the parallel registers data are copied into a serial output register and shifted out serially.

Fig. 2 shows the chip size normalized to capacity for some commercial CCD memories. The upper curve is for serpentine and the lower for SPS organizations. The normalized size decreases with increasing capacity because CCD cell size gets smaller and smaller in each new product. Independent of capacity the normalized size of the serpentine organization is found to be twice the size of SPS organization. Consequently, SPS is preferable when a high capacity is to be realized.

To estimate the realistic capacity we note, that a typical cell size is about 400 μm^2 when conventional layout rules are used. With an overhead for wiring, refresh amplifiers, input and output buffers, and clock circuits of 50 %, an effective cell size of 600 μm^2 results. For these values, the chip size would amount to 9.83 mm^2, 19.7 mm^2, and 39.3 mm^2 for capacities of 16k, 32k, and 64k respectively. Consequently a 32k capacity will be feasible without reducing cell dimensions.

Advanced means to reach higher packing density such as electrode per bit organization, offset gate technology, and multi-level storage were studied. But they were not used in order to reach reliable operation with standard technology. There was no attempt to reach very high clock frequencies. A simple calculation shows why this is not necessary in large systems when power dissipation is to be minimized. The clock driver power dissipation in a system comprising m CCD chips is

$$N_{V1} = m \, N_{V,B} = m \, c_1 \, f_1 \tag{1}$$

where $N_{V,B}$ is the clock driver power for one chip, c_1 is a constant and f_1 is the clock frequency. If n chips are multiplexed to one data line, the clock frequency can be lowered to $f_n = f_1/n$. The clock power is now

$$N_{Vn} = m \, N_{V,B} + N_{V,M} = m \, c_1 \, f_n + m \, N_{V,Mux} \tag{2}$$

where $N_{V,M}$ is the total multiplexer power, and $N_{V,Mux}$ is the power dissipation in one multiplexer stage. From (1) and (2) the ratio of power dissipation is

$$\frac{N_{Vn}}{N_{V1}} = \frac{1}{n} + \frac{N_{V,Mux}}{c_1 \, f_1} \tag{3}$$

If $n > 10$ it follows that the power dissipation in the multiplexed system is lower if the power dissipation in one multiplexer stage is lower than the clock power in one CCD chip at clock frequency f_1. As a consequence of this result no attempt was made to maximize the upper clock frequency.

4. Comparison of different organizations

Advantages and disadvantages of different organizations (like serpentine, SPS, LARAM) have been discussed in the literature /3/, /4/, /5/. These considerations will be detailed here to find a suitable organization for the 32 kbit chip. This will be based on very rough approximations. However a quantitative analysis has been made which gave virtually the same results.

The results are compiled in Fig. 3. The table is based on the following assumptions.
(1) The maximum clock frequency is limited by refresh amplifiers and output buffers.
(2) The minimum clock frequency is given by the maximum storage time between refreshs, which in turn is mainly determined by dark currents. i is the number of independently

accessible blocks, n_p is the length of the parallel registers and n_s is the length of the serial registers for the SPS organization. In the cases of serpentine and LARAM organizations these parameters apply in a similar manner. The following conclusions can be drawn from the table. In some SPS-organized versions the minimum clock frequency is not lower than the maximum. Those organizations are not useful since no flexibility is provided. In the serpentine arrangement low minimum clock frequencies are achievable because there are few storage locations between refresh amplifiers in this structure. This means on the other hand a large chip area. The chip areas for 32 kbit are compared in Fig. 4. Obviously a small number of individually accessible blocks results in a smaller chip area. The SPS organization results in the least area of the alternatives compared. Furthermore the chip area is nearly independent of the parameters n_p and n_s, so other criteria can be used to optimize n_p. Consequently the SPS structure is preferable for the given application. The price to be paid for higher density is a smaller span between maximum and minimum clock frequencies.

5. Choice of SPS blocks

The remaining task is now to partition the 32 kbit into SPS blocks. From test circuits it is known that problems with dark currents are most likely /6/. Therefore a tradeoff between block size and tolerance to dark currents has to be made. This is because larger blocks give better overall packing density but are more susceptible to dark currents and have a higher minimum clock frequency thus decreasing the useful range between maximum and minimum clock frequencies. Several alternatives are compared in Fig. 5.

Case 1 assumes 32 1kbit blocks, one of which can be selected by means of a 5 bit address. Smaller blocks in case 2 lead to a smaller latency time and lower minimum clock frequency. Case 5 assumes 4 blocks of 1kbit each with individual refresh amplifiers, which are connected in a ring structure. But only access to a 4kbit ring is provided. Fig. 6 shows the size of the CCD array for 32 kbit as a function of block size a, b and the width of wiring channels between blocks c and d. The resulting chip areas show that only 4kbit blocks lead to an acceptable chip size.

6. Design of SPS blocks

The SPS blocks should be as small as possible. In test circuits we found that 64 bit long registers with an inverter as output amplifier can be made with channel widths down to $w_s = 1.5$ μm. This means, that the channel width is only limited by the resolution of optical lithography. The electrode length l on the mask was fixed at 9 μm as is shown in Fig. 7, the interelectrode gap d is 5 μm, resulting in a cell length L of 28 μm. The length of the cell determines the period of the parallel registers. Fig. 7 shows the layout of the transfer region between the serial input register and the parallel registers. $\emptyset_{S1}$ and $\emptyset_{S2}$ are the serial clocks, $\emptyset_{P1}$ and $\emptyset_{P2}$ the parallel clocks. $\emptyset_T$ is the transfer clock. The distance d_p of the parallel registers can be decreased if parallel registers are multiplexed with one cell of the serial register. This means that there are several transfers from serial to parallel registers during one clock period of the parallel clocks. In the lower part of Fig. 7 the layout of a multiplexed SPS block is shown. The period L is reduced to one-half the cell length, which gives 14 μm and the channel width w_p is fixed at 6 μm. The resulting distance between channels d_p is 8 μm, giving a cell size of 392 μm^2. The order of multiplexing cannot be greater than the number of storage electrodes per serial cell. For the two-phase system used here the order of multiplexing is two.

Fig. 8 shows the timing diagram for multiplexing. $\emptyset_{S1}$ and $\emptyset_{S2}$ are the clocks for the serial input registers. $\emptyset_1$ and $\emptyset_2$ are for the serial output register and are the only clocks which just be provided from outside the chip. The transfer clocks $\emptyset_{T1}$ and $\emptyset_{T2}$ for transfer between serial and parallel registers and the parallel clocks $\emptyset_{P1}$ and $\emptyset_{P2}$ are generated on the chip to simplify the application in memory systems. As argued above, 4kbit SPS blocks are necessary to realize an acceptable chip size. The length of

parallel and serial registers can be determined for minimum clock power. Total clock power is the sum of power in the serial and the parallel registers

$$N = N_P + N_S \tag{4}$$

Clock power is proportional to register length n_S, n_P, oxide capacitance C, and to the clock frequency f_C. Thus, total clock power is

$$N = C U^2 f_C (2n_S + n_P) \tag{5}$$

Minimum clock power results if

$$2n_S = n_P \tag{6}$$

is chosen.

One important parameter is the transfer inefficiency ϵ . Since in each transfer charge loss occurs, the maximum number of transfers between refreshs is limited. Because of the sensitivity of refresh amplifiers, tolerances in technology and temperature dependence, a total charge loss of 10 percent is acceptable. This means that the product of transfer inefficiency ϵ and register length n is bounded by

$$n \epsilon < 0.1 \tag{7}$$

At a clock frequency of 2 MHz transfer inefficiency is $5 \cdot 10^{-4}$ as measured in our test circuits. Consequently no more than 200 transfers are possible. Thus, for a two-phase CCD, the maximum register length is 100 bit. Since the transfer inefficiency increases with increasing frequency, charge loss occurs mainly in the fast-clocked serial registers. To minimize charge loss this should be short. A value $n_S = 32$ was determined as a compromise between low charge loss and low clock power. The length of the parallel registers would be 128 bit. Because multiplexing is used, each parallel register is shortened by the multiplex order. For two phase operation (multiplex order equal two) the resulting length is 64 bit. This in turn results in a block with minimum power and maximum packing density.

Fig. 9 shows a microphotograph of a 4kbit SPS block, as it has been realized in n-channel two layer silicon-gate technology with local field oxidation. Size of the block is 1.74 mm^2.

7. Technology problems

Dark currents are responsible for the minimum clock frequency. In a 64-bit test delay line storage time was up to 100 ms, corresponding to 650 Hz minimum clock frequency. Because measuring the minimum clock frequency is time consuming, life-time measurements using large (10^5 to 10^8 μm^2) MOS-varactors were used for process characterization. We found that for 2 ms refresh intervals lifetimes between 40 and 80 s are sufficient. Similar results were found using a guarded diode structure after Grove and Fitzgerald /8/. This measurement is difficult, because currents in the 10^{-12} A range must be measured. Measurements using pulsed varactors /9/ are much easier to perform. One major problem is in local generation of dark current at generation centers. Using a CCD register the centers can be located. Fig. 10 shows the contents of 64 bit register at 50°C after 25 ms integration time. The different amplitudes are caused by different dark currents in the corresponding cells. Etching for instance with an etching solution after Secco /10/ showed a pronounced correlation between stacking faults and local dark current generation. We found that a density of 3 per 10^6 μm^2 is acceptable. Process optimization for low fault density has been done.

8. Design of a 32kbit memory chip

Fig. 11 shows the block diagram of the 32kbit chip. Beside the 8x4kbit memory blocks with associated circuitry, the clock generator, reference voltage generator and data input/output buffer are essential parts. The clock generator drives six clock drivers

which are associated with the individual SPS blocks. This leads to a better layout than would one clock driver for all blocks together.

The reference voltage generator consists of a CCD register which is similar to a 4kbit block, a refresh amplifier and a feedback circuit which generates a voltage corresponding to the middle value of the voltages produced by a logical zero and a none. Thus, the reference level is always optimum regardless of temperature, frequency, and technology tolerances.

All other circuitry which is necessary for operation of the individual blocks is associated directly with the blocks. Besides the clock drivers CD these are: refresh amplifier RA, and selection circuits SC. One of eight blocks is selected for reading or writing by means of a 3 bit address, while in all other blocks information is synchronously recirculated. The selection circuits also provide for recirculation of the stored information as long as no write operation takes place.

The 32kbit chip is currently being laid out, so no measured data can be given. All essential parts like 4kbit blocks, refresh amplifiers and reference voltage generator have been realized on test circuits. From these results, expected data can be given as shown in Fig. 12. Chip size will be 24 mm^2. For the 4 k times 8 organization the average latency time is 1 ms at 2 MHz clock frequency. Clock frequencies will be between 0.5 - 2 MHz for a temperature range from 0 to 50°C. Two external clocks are necessary, and load capacitance is only 10 pF. Supply voltages are +12 V and -5 V, and power dissipation is 230 mW at 2 MHz clock frequency.

8. Conclusion

Considerations for the design of digital CCD memory chips have been discussed. The principal aim of this work was to design a simple, reliable memory chip which would give good yield and therfore low cost. It has been shown, that with conventional technology and layout rules a 32kbit chip is economically feasible. To achieve 64kbit capacity some non-standard measures have to be taken resulting in a near-term economic disadvantage.

Acknowledgments

References

/1/   A. Ludloff, M. Minker, 1971, "Theorie und Praxis des Signalprozessors von Rundsichtradars", Wiss. Ber. AEG-Telefunken 45, pp. 9-17

/2/   A. Ludloff, 1976, "Zum Stand der Signalverarbeitung bei Rundsichtradars", Nachrichtentechn. Zeitschrift 29, pp. 647-651

/3/   L.M. Terman, L.G. Heller, 1976, "Overview of CCD memory", IEEE SC-11, pp. 4-10

/4/   S.D. Rosenbaum et.al., 1976, "A 16384-bit high-density CCD memory", IEEE SC-11, pp. 33-40

/5/   C.H. Sequin, M.F. Tompsett, 1975, "Charge transfer devices", Supplement 8 to Advances in electronics and electron physics, Academic Press

/6/   D.F. Barbe, 1975, "Imaging devices using the charge coupled concept", Proc. IEEE 63, pp. 38-67

/7/   J.M. Chambers et.al., 1974, "CCDs as drum and disc equivalents", 1974 WESCON, Professional program, Los Angeles

/8/   A.S. Grove, D.J. Fitzgerald, 1966, "Surface effects on pn-junctions", Sol. St. Electronics 9, pp. 783-806

/9/   A.F. Tasch et.al., 1973, "Dark current and storage time considerations", CCD Applications Conference, San Diego

/10/   F. Secco d'Aragona, 1972, "Dislocation etch for $\langle 100 \rangle$ planes in silicon" J. Electrochem. Soc. 119, pp. 948-951

DISCUSSION

J.J.Stapleton

When you say low cost, is it less then 100 dollars a chip?   When will it be 1 dollar?

Author's Reply

In R and D we cannot talk dollars!
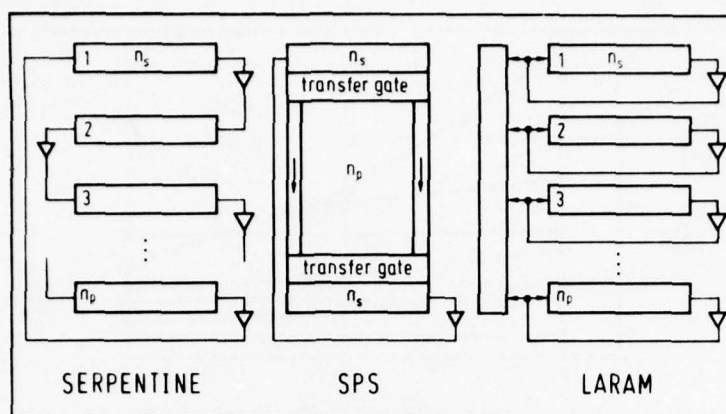
Fig. 1　Comparison of different organizations
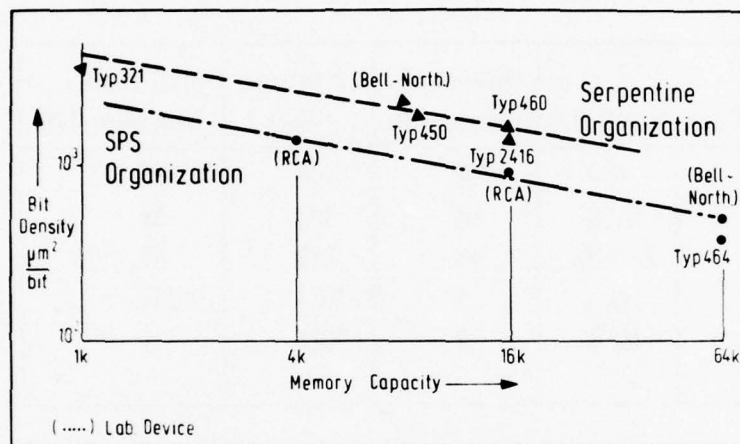
Fig. 2　Bit density for some CCD memories

i	n_p	n_S	pin count		clock frequency/MHz			
					max.		min.	
			Serp.	SPS	Serp.	SPS	Serp.	SPS
1	64	512	10	12	10	10	0.13	8
2	64	256	12	14	10	10	0.13	4
4	64	128	16	18	10	10	0.13	2
1	128	256	10	12	8	10	0.25	8
2	128	128	12	14	8	10	0.25	4
4	128	64	16	18	8	10	0.25	2
1	256	128	10	12	5	8	0.51	8
2	256	64	12	14	5	8	0.51	4
4	256	32	16	18	5	8	0.51	2
4	64	128	12	14	10	10	0.13	2
4	128	64	12	14	8	10	0.25	2
4	256	32	12	14	5	8	0.51	2

Fig. 3　Comparison of different CCD organizations
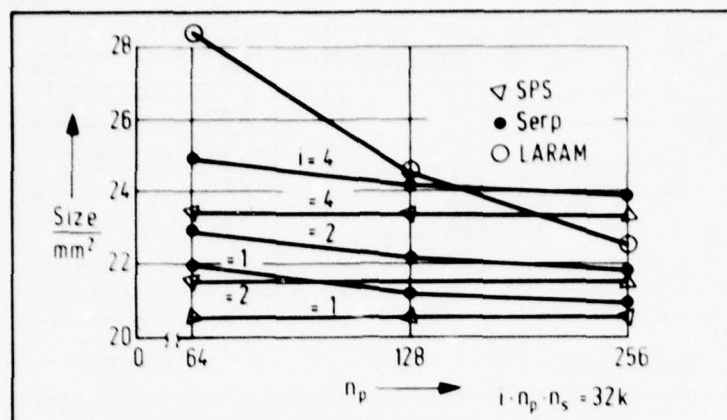
Fig. 4   Chip size for 32 kbit CCD

Case	Block size	Number of refresh amp.	Latency/µs (2 MHz)	min. clock frequency /kHz
1	1k	32	256	125
2	0.5k	64	128	63
3	2x 0.5k	64	256	63
4	4k	8	1024	500
5	4x 1k	32	1024	125

Fig. 5   Comparison of different SPS organizations

Case	Size of CCD array	F / mm^2 c = 0.15mm	c = 0.05mm
1	(8b + 9c)(4a + 5c)	25.7	19.3
2	(8b + 9c)(4a + 9c)	29.2	20.2
3	(8b + 9c)(4 (2a + d) + 5c)	26.9	20.2
4	(8b + 3c)(4a + 5c)	18.9	17.6
5	(4 (2b + d) + 5c)(2 (2a + d) + 3c)	22.8	19.3

Fig. 6   Size for different SPS organizations
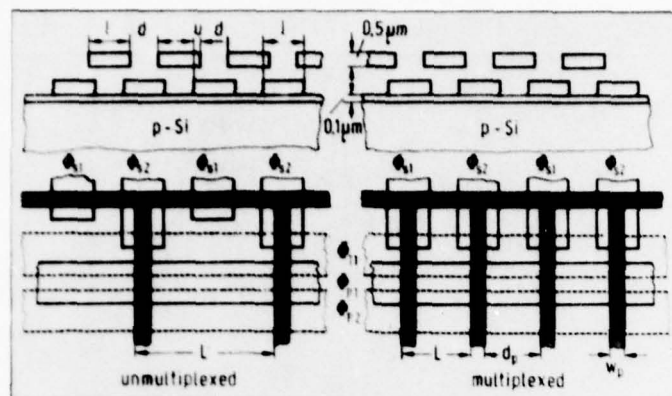
Fig. 7   SPS electrode and layout configuration

Φ_{S1}

Φ_{S2}

$10V$

Φ_{T1}

Φ_{P1}

$1\mu s$
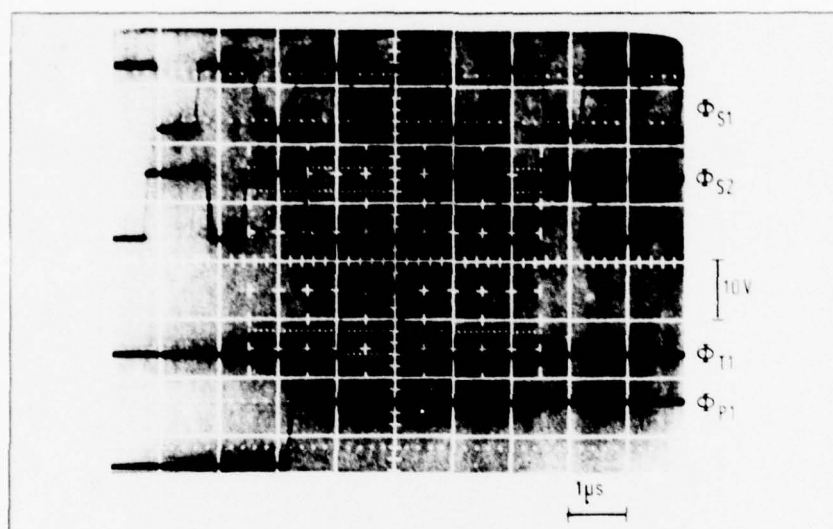
Fig. 8   Clock wave forms for the SPS array

0.93mm

187mm
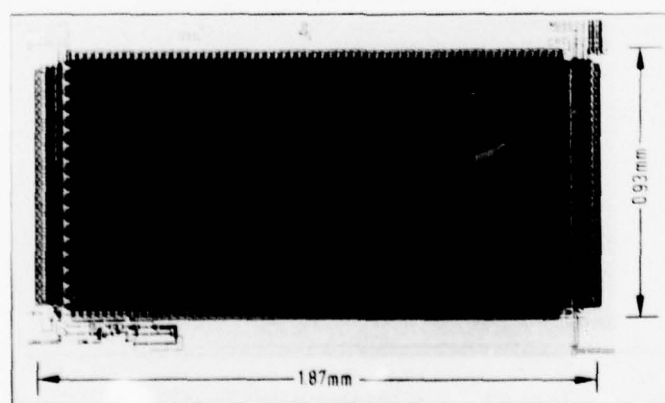
Fig. 9   4 kbit SPS array
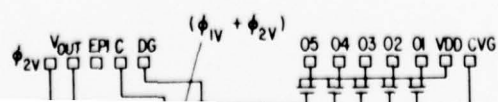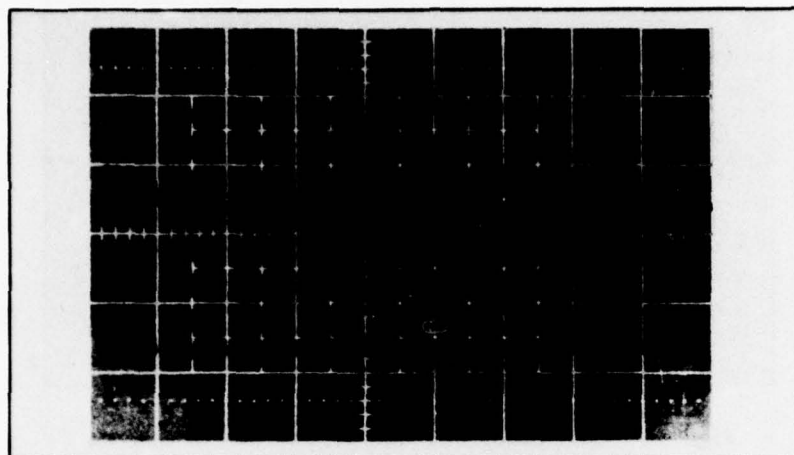
Fig. 10   Output due to dark current
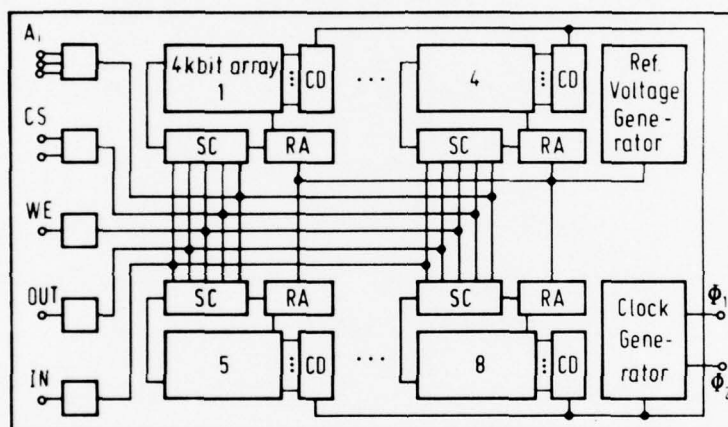
Fig. 11   Block diagram of 32 kbit memory chip

Organization:	8 x 4096 bit
Average access time:	1 ms ... 4 ms
Two overlapping clocks:	0.5 MHz ... 2.0 MHz
Clock capacitance:	10 pF
Output :	open drain, max. 4 mA
Power supplies:	12 V, -5 V
Temperature range:	0 - 50 C
Power dissipation:	230 mW (2 MHz)

Fig. 12   Preliminary specifications

ELECTRO-OPTICAL PROCESSING

OF SIGNALS & IMAGES

D. Roy
M.A. Copeland
C.H. Chan

Carleton University
Ottawa
Canada

INTRODUCTION

The use of an LED/CCD structure to perform transversal filtering of electrical signals has been previously demonstrated (1, 2). Basically, the action of convolution is achieved by progressive summation as against the simultaneous summation of tap weighted signal samples as in the split electrode CCD filter. Summation is done in a given CCD potential well as the well is first created and then progressively clocked to the output end of the device. The multiplications necessary in a convolution operation are obtained by means of an optical mask which controls the amount of light from a signal modulated LED reaching the various CCD electrodes. Although the arrangement is mask programmable, the real merit of this principle probably lies in the area of image processing where a focussed image naturally provides a space varying illumination and where the coefficients of a desired transformation can be applied as a time varying signal. This leads to an electrically programmable image processor where the processing is done functionally inside a CCD chip.

This paper will describe the two types of processing with experimental results and discussions on performance limitations.

II. EXPERIMENTAL ARRANGEMENTS & RESULTS

The basis of electro-optical processing is the sensitivity of a charge coupled device to light. Both types of linear processing mentioned in the previous section rely on being able to modulate the charge gathering capability of the CCD delay cells. The important contrast between these two approaches is in the way the coefficients of a desired transformation are applied. In the case of electrical signal processing, the signal samples move past the weighting coefficients held on a transmission mask and in the case of image processing the reverse is true. The following two sections are devoted to discussions of preliminary results on both types of processing.

(a) Processing of electrical signals

A simplified block diagram of the experimental arrangements of the processor for processing of electrical signals are shown in Fig. 1. The input analog signal modulates the output of a light emitting diode. Since the amount of photogenerated carriers at the silicon surface is proportional to both the intensity of light reaching the silicon surface and the time the light source is on. the LED can be either intensity modulated or duty cycle modulated. Both modulation schemes can be employed to obtain premultiplication of the input signal as is required in some applications (3). The coefficients of a desired filter function are encoded onto a glass mask by varying the ratio of the transparent to opaque area. The LED is turned on once every CCD clock period when there are valid potential wells under the CCD electrodes and the result of a convolution is directly available at the output of the device. Since charge is detected only once at the final output of the device, performance comparable to a linear imager is expected.

To achieve tap weights of both polarities multiplexing is necessary since only one polarity of charge can be stored in the potential wells of a CCD. Figure 2 (a) and (b) respectively show the circuit arrangements for time and space multiplexing. In the time multiplexed scheme two adjacent charge storage sites are used to represent one tap coefficient. Every alternate delay cell represents a negative charge storage site and the output of the filter is alternately switched between the positive and the negative input terminal of a differential amplifier. A simple band pass filter (truncated sine-wave in time domain, Fig. 3) was realized using time multiplexing. This approach in general gives rise to very fast changing mask patterns which are more susceptible to the tap weight errors that arise from the lack of collimation of the incident light beam and finite distance of separation between the mask and the actual device. An additional limitation of this approach may arise from the requirement of 2N CCD delay elements to realize an N-tap filter and the consequent reduction of bandwidth capability by one-half. A more practical approach appears to be the space multiplexed scheme where the signal is divided in space between two CCD line imagers. The coefficient masks are so designed that subtraction between the outputs from the two CCD channels give rise to the desired filter response. Figure 4 shows a demonstration of this scheme for a 91-tap low pass filter. The actual photomask employed in the experiment is shown in Fig. 4 (a) and the observed impulse response is shown in Fig. 4 (b). As an additional demonstration of this concept the experimental and predicted frequency response of a 91-tap Hamming weighted band pass filter is shown in Fig. 5. In each of the above experiments the coefficient mask was designed by applying a suitable bias to the given impulse response as is evident from Fig. 4 (a). This approach, although it reduces the dynamic range by one-half has the advantage of design simplicity. Further a number of filter functions can be made on an area array imager with a common bias channel. The Fairchild's 256 interleaved photoelement linear imager is particularly suitable for realizing a general filter function using the space multiplexing approach. All results reported in this paper were obtained using simple CCD delay lines which were not specifically designed for imaging applications. Two separate CCD chips were used to

achieve space multiplexing.

(b) Processing of images

The experimental arrangements for image processing consist of a coefficient memory where the coefficients of a desired transformation are stored in digital words or, can be generated using algorithms. These coefficients are sequentially clocked out to modulate either the light intensity of the LED or the period of the CCD clocks. For processing photographic negatives the former method can be used. The latter method would allow direct image processing.

In the demonstration experiment discussed here an optical mask was used to define a linear image. The cosine and the sine coefficients of a 32-point Discrete Fourier Transform were stored in a bank of PROM's. All coefficients were stored as positive numbers, implying a bias applied to the coefficients. This bias is reflected at the output as a fixed offset and hence can be subtracted out in a differential amplifier. The experimental results of correlation of the cosine coefficients with a rectangular image pattern of Fig. 6 (a) is shown in Fig. 6 (b). The components of the cosine transform are obtained every 32nd sample at the output of the CCD as shown in Fig. 6 (c). The results on this simple experiment indicate feasibility of processing fixed images directly inside a CCD chip. The concept can be extended to CCD imaging array for obtaining two-dimensional linear and separable transforms (e.g., Hadamard, Fourier, etc.) (4). The area array imager will provide transforms of all the columns at the same speed as that for a single column. To obtain the transforms of the rows digital/analog techniques (5,6) will have to be used at the output of the CCD. The concept of linear image processing can be applied for processing of one-dimensional variables e.g., monitoring of fluid level or diameter of a wire for industrial control.

III. PERFORMANCE CONSIDERATIONS

The performance of the electro-optical processor is limited by accuracy, linearity and noise. The inherent limitations of a CCD imager affect the performance of both types of processing discussed earlier. The discussions made in this section are primarily in the context of electrical signal processing, however they apply to image processing applications as well. The experimental vehicle was a Blackman window filter function (7) which provides a peak side lobe amplitude of about -57dB. The window function has all positive taps hence can be implemented using single CCD channel as for processing of linear images.

Figure 7 shows the observed and predicted frequency response of a 32-tap Blackman window function. The theoretical response takes into account the error of coefficient quantization involved in the preparation of the photomask. The relatively wide channels of the devices used provided a quantization step size equivalent to $12\frac{1}{2}$ bits of accuracy in a digital system. The observed errors in the stop band region are mainly due to errors associated with the variations of the device parameters. The tap weight errors introduced by lack of collimation of the incident light beam are very small for a smoothly varying mask pattern. For the filter function under consideration the tap weight errors from this source will introduce slight roll off of the frequency response close to the Nyquist frequency limit. This error can further be reduced by using a suitable lens. No lens was used in any of the experiments reported here. Due to optical injection of charge major source of tap weight errors are expected to arise from any variation of transmittance of light through the multilayer electrode structure of a charge coupled device (2-phase polysilicon gate CCDs were used in the experiments). The transmittance of light through the electrodes of a CCD is dependent on the thickness of each electrode layer (8). Hence any variation of thickness of any of the layers along the CCD channel will cause tap weight errors. A computer-aided sensitivity analysis was done by recording the transmittance variations by perturbing the values of layer thicknesses from their nominal values. It was observed that the polysilicon layer particularly, produces a wide variation of transmittance even for a small change in thickness around the nominal value. For the devices used the red LED light (0.66mm wavelength) was found to be less sensitive to the changes in the polysilicon layer thickness than yellow and green light. The experimental results however do not indicate any significant variation of light transmittance along the CCD channel. The observed frequency response of the Blackman window function shown in Fig. 7 indicates tap weight inaccuracies better than 0.5% (as obtained by Monte Carlo simulation (9)). Although acceptable accuracies have been obtained for other filter functions e.g., a 91-tap Hamming window function (1), variation of light transmittance along the CCD channel is a potential source of tap weight errors which must be recognized. The sensitivity of light transmittance to layer thickness variations may be greatly reduced by using special electrode structures (10).

Additional tap weight errors may arise from lateral diffusion of carriers generated in the neutral bulk beyond the depletion layer edge. In the context of CCD imagers the effect of lateral diffusion is commonly expressed in terms of a modulation transfer function (MTF) (11). The resolution degradation due to lateral diffusion is spatial frequency dependent. Thus the amount of error introduced is dependent on the nature of the tap weight mask. For fast changing tap patterns i.e., for filters with passband close to the electrical Nyquist frequency limit the effect of lateral diffusion will be more pronounced because of high spatial frequency content of the mask pattern. The tap weight errors due to lateral diffusion introduce a roll-off in the frequency response particularly in the high frequency region. Using red LED light the computed MTF behaviour of the device suggests that this error is only significant (>1%) for filters having passband extending beyond 1/4th the CCD clock frequency. Increasing the depletion layer width will confine the generation of most of the carriers to within the depletion layer region, thereby reducing the effect of lateral diffusion. Lower wavelength of the incident light e.g., green light will produce a similar result by reducing the penetration depth.

The sources of tap weight errors discussed above in the context of an electrical filter, will cause degradation of the image reproduced by a CCD imager. Consequently an image slightly different from the one focussed onto the device will be processed by the image processor. The coefficient errors of the image processor arise from quantization of the coefficients in a limited number of digital bits. The signal to noise ratio due to such quantization increases approximately 6dB for each bit added.

Since negative coefficients are handled by applying a bias to all the coefficients, there is a loss of resolution of one-bit for a given number of bits/word used.

The measured harmonic distortion at the output of the Hamming weighted bandpass filter (see Fig. 5) is shown in Fig. 8. The linearity of processor is dependent on the linearity of the input stage where light power is converted to charge and linearity of the output stage where charge is converted to an equivalent voltage. The amount of charge stored at the surface depletion layer of a CCD over a given integration time is directly dependent on the light intensity and very weakly dependent on the depletion layer width. The input non linearity is only significant when the depletion layer width changes by a wide margin during the on time of the LED. Normally, this change is very small even from no charge to full charge condition of a potential well. Consequently, the observed non linearity is mainly due to the charge detection stage. Analysis of the harmonic distortion of the output stage alone predicts a second harmonic distortion within 3 to 4 dB of the measured second harmonic distortion. Thus linearization of the output stage using on-chip amplifiers will further improve the linearity of the processor.

Noise measurement is important to the maximum speed of operation of the processor. This is because increasing the speed of operation reduces the integration time i.e., the on time of the LED thereby lowering the signal to noise ratio. The measurement of noise was performed on the Blackman window function using correlated double sampling at the output. The results reported here however do not indicate optimum noise performance, since the devices used were old and exhibited poor surface states effect. With 24 µsec integration time, the measured rms signal to noise ratio for peak signal was 62 dB. The luminous intensity of LED used was around 3 mcd. Reducing the integration time by a factor of four, the above signal to noise ratio decreases to about 50 dB. Lowering the integration time however increases the linearity of the processor through reduction of the bias charge. The noise sources associated with the processor are identical to that of a CCD imager. However, the shot noise associated with the optical injection of charge is filtered at the output by the filter function. Contrary to the split electrode filter, the dark current noise and the surface state noise are independent of the filter transfer function as these noises are introduced following the coefficient multiplications. An additional noise component may arise due to jitter in the LED drive clock. Any random variation of the width or the amplitude of the LED drive clock from sample to sample will introduce white noise at the output. The output signal to noise ratio due to this noise component is independent of the filter transfer function consequently tight clock tolerance is necessary to limit this noise component.

The use of an area-array CCD imager for processing of images introduces additional distortions due to non-uniformities across the time-multiplexed channels of transfer inefficiencies, gains, leakages and offset levels (12). These distortions can be significant unless good devices are used.

IV. CONCLUSIONS

This paper discusses an electro-optical approach for processing of electrical signals and optical images. Initial results presented on both types of processing encourage further work particularly, in the area of image processing. The important performance limitations have been discussed in the context of technology limitations of charge coupled devices. The effect of major sources of limitations can be reduced with improved design of CCD imagers.

REFERENCES

(1) "An optical CCD convolver", M.A. Copeland, D. Roy, J.D.E. Beynon, & F.Y.K. Dea, IEEE J. Solid State Circuits, Feb. 1976, pp. 84-87.

(2) "Electro-optical processing of signals and images using CCD's", D. Roy, M.A. Copeland, Proc. IEDM, Dec. 1976, pp. 239-242.

(3) "The Chirp-Z transform algorithm", L.R. Rabiner, R.W. Schafer, & C.M. Rader, IEEE Trans. Audio Electroacoust., vol. 17, pp. 86-92, June 1969.

(4) "Picture processing and digital filtering", T.S. Huang (Ed.), Springer-verlag, Berlin, 1975.

(5) "An electronically programmable transversal filter", Y. Haque, & M.A. Copeland, Proc. IEDM 1976, pp. 27-30.

(6) "Image contour extraction with analog MOS circuit techniques", P.I. Suciu, & D.A. Hodges, IEEE. J. Solid State Circuits, vol. 12, No. 1, Feb. 1977, pp. 65-72.

(7) "Digital signal processing", A.V. Oppenheim, & R.W. Schafer, Prentice-Hall, 1975.

(8) "Transmittance of air/SiO$_2$/Polysilicon/SiO$_2$/Si structures", C. Anagnostopoulos, & S. Sadasiv, IEEE J. Solid state circuits, vol. 10, No. 3, June 1975, pp. 177-179.

(9) "The design and operation of practical charge-transfer transversal filters", R.D. Baertsch, W.E. Ergeler, H.S. Goldberg, C.M. Puckette, & J.J. Tiemann, IEEE J. Solid state circuits, vol. 11, No. 1, Feb. 1976, pp. 65-74.

(10) "Transparent metal oxide electrode for CID imager", D.M. Brown, M. Ghezzo, & M. Garfinkel, Feb. 1976, pp. 128-132.

(11) "Carrier diffusion degradation of modulation transfer function in charge coupled devices", D. Seib, IEEE Trans. Electron devices, vol. 21, no. 3, March 1974, pp. 210-217

(12) "Distortion effects due to multiplexing of charge transfer structures", D. Roy, C.F. Rahim, & M.A. Copeland, IEEE Trans. Electron devices, vol. 24, no. 6, June 1977, pp. 679-684.

RELATED MATERIAL

1. I. Lagnado and H.J. Whitehouse, "Signal Processing Image Sensor Using Charge-Coupled Devices",
    Proc. Int. Conf. on Technology and Applications of Charge-Coupled Devices, Edinburgh, Scotland,
    September 1974, pp. 198-205.

2. R.P. Bocker, "Matrix Multiplication Using Incoherent Optical Techniques", Appl.Opt., Vol. 13,
    July 1974, pp. 1670-1676.

3. K. Bromley, "An Optical Incoherent Correlator", Opt. Acta., Vol. 21, No. 1, 1974, pp. 35-41.
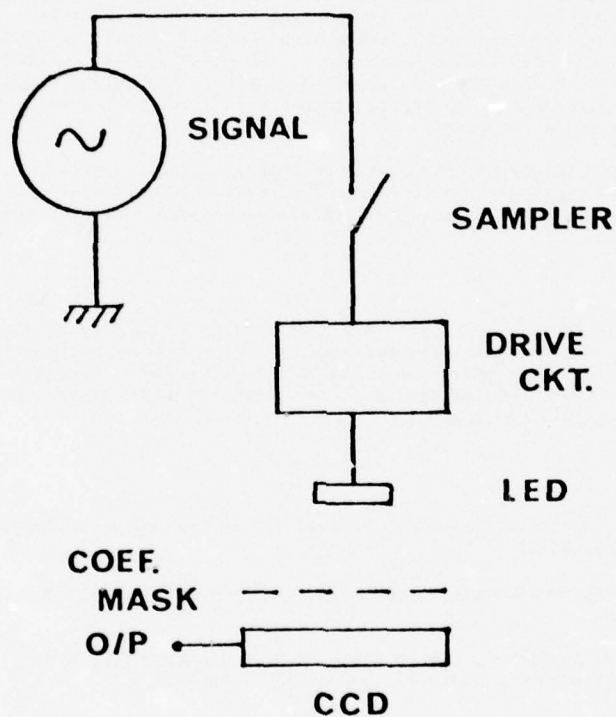
Fig. 1   A block diagram of the experimental set up for signal processing.

1 BIT DELAY
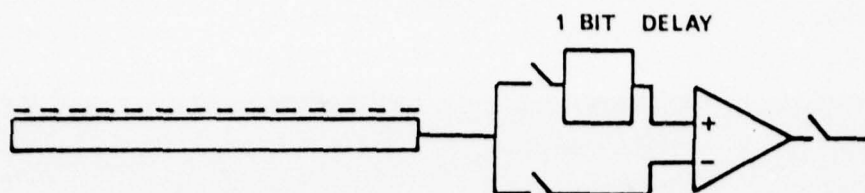
(a) TIME MULTIPLEXING
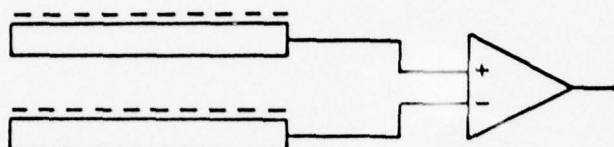
(b) SPACE MULTIPLEXING

Fig. 2   Two approaches for realizing positive and negative tap weights.

two adjacent storage sites

(a)

Response (dB)

— Theory
--- Experiment

Signal Freq. / Clock Freq.

(b)
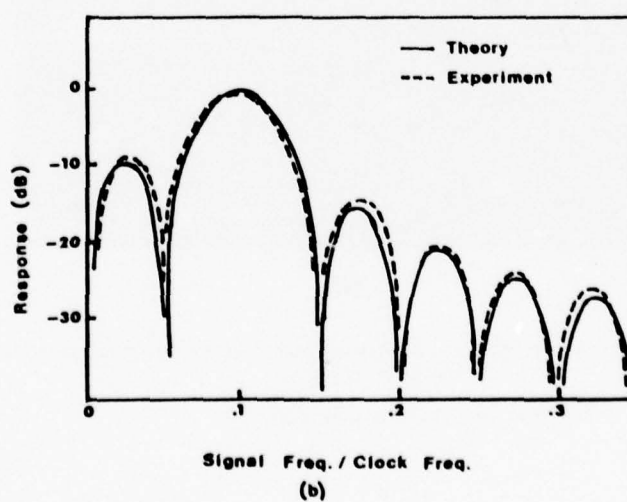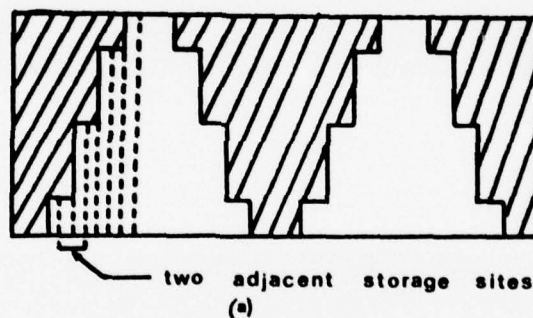
Fig. 3   (a)   The mask pattern for a 21-tap two cycle sine wave filter exposing 42 CCD delay elements.

(b)   The observed and predicted frequency response.
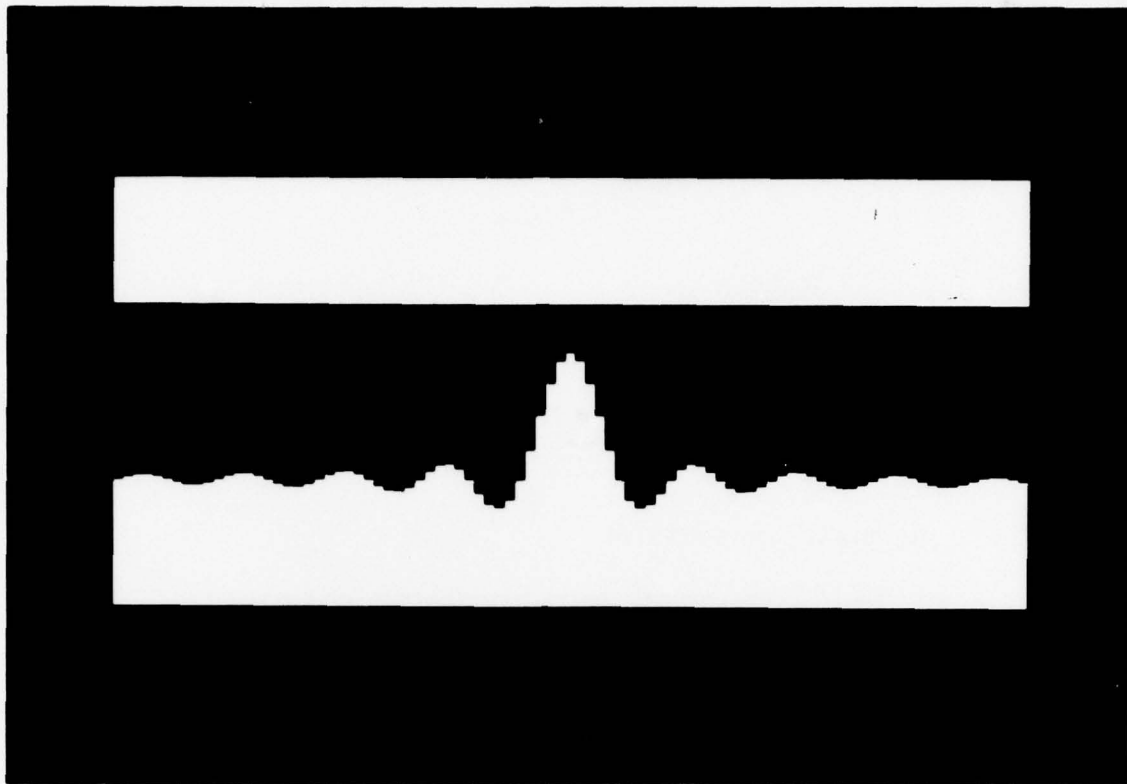
(a)

(b)

Fig. 4   (a) A photograph showing the designed mask pattern for a
              91-tap low pass filter.

        (b) The observed impulse response.
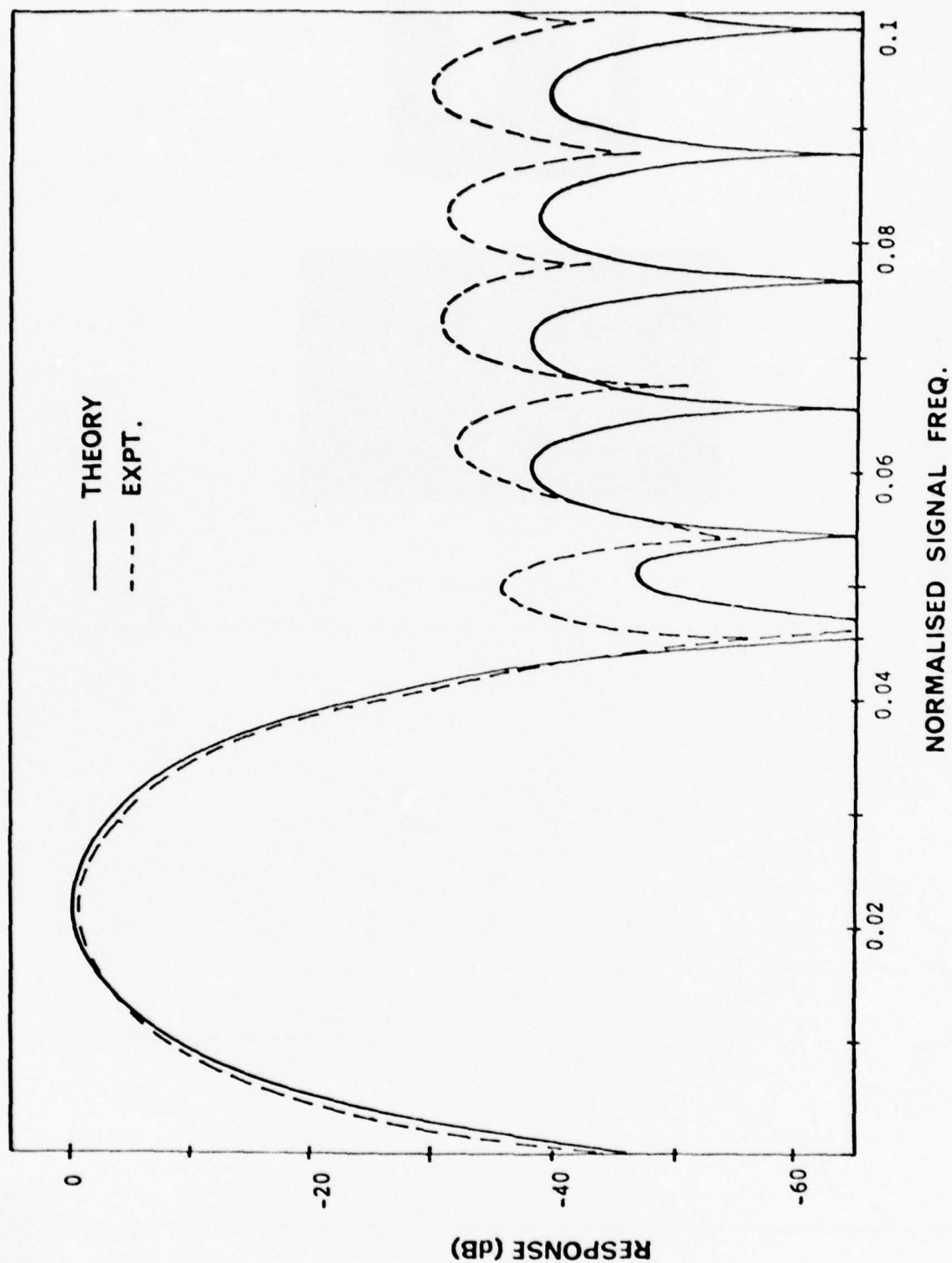
Fig. 5   The observed and the theoretical frequency response of a Hamming weighted bandpass filter.

(a)

(b)

input cosine coefficients

(c)
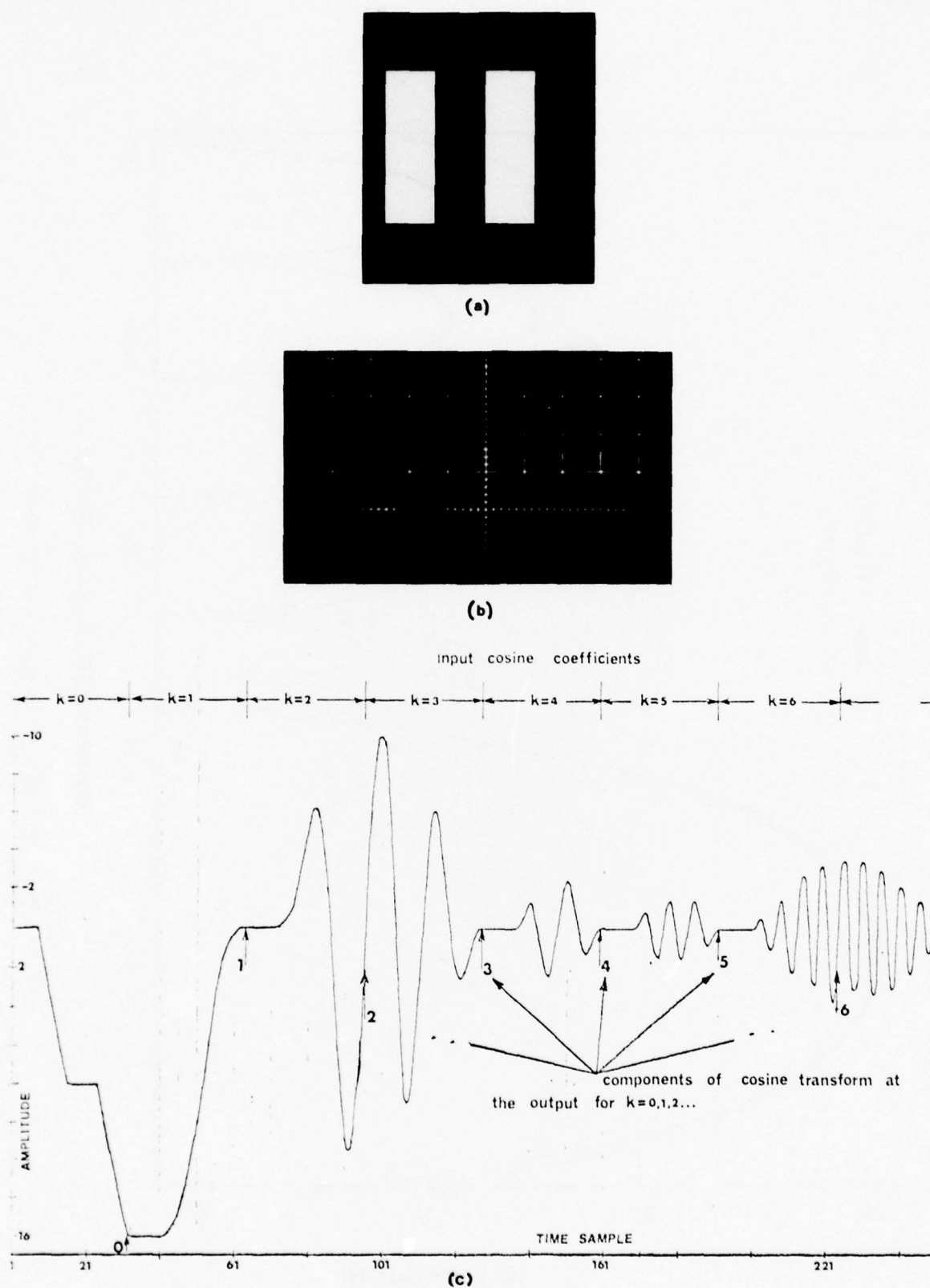
Fig. 6 (a) A photograph showing a square wave image pattern used
            in the image processing experiment.

        (b) The experimental results of correlation of the image
            with the cosine coefficients of a 32-point DFT.

        (c) The theoretical results.
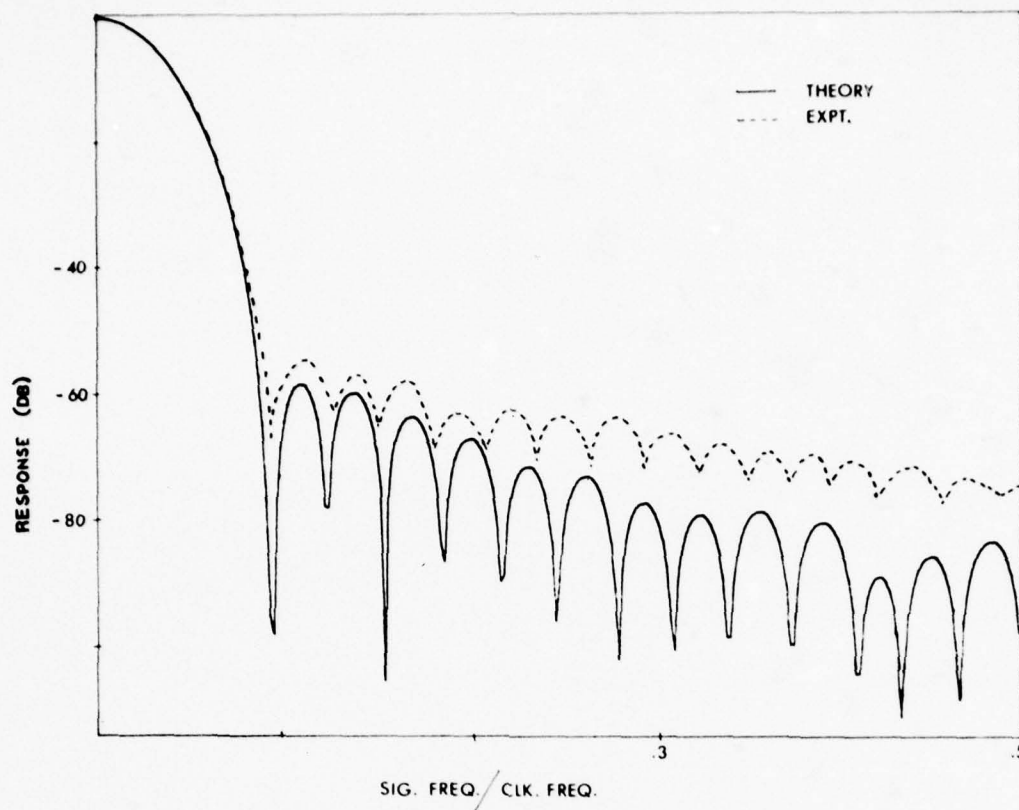
Fig. 7   The observed and predicted response of a 32-point Blackman
window function.   A peak side lobe amplitude of ~56 dB
has been achieved.
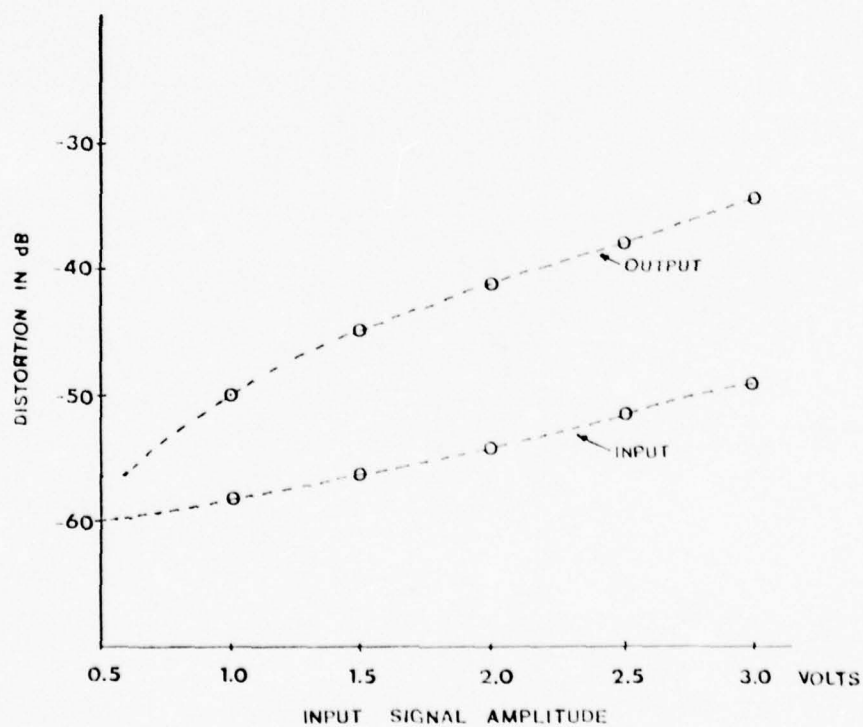
Fig. 8   The measured harmonic distortion of the output signal and
the LED drive current.

CHARGE INJECTION DEVICE (CID)
HADAMARD FOCAL PLANE PROCESSOR
FOR IMAGE BANDWIDTH COMPRESSION*

G.J. Michon, H.K. Burke, T.L. Vogelsong

GENERAL ELECTRIC COMPANY
CORPORATE RESEARCH AND DEVELOPMENT
P.O. Box 8
Schenectady, New York 12301

and

P.A. Merola

GENERAL ELECTRIC COMPANY
AEROSPACE ELECTRONIC SYSTEMS PRODUCTS DEPARTMENT
831 Broad Street
Utica, New York 13503

ABSTRACT

A Charge Injection Device (CID) solid-state video sensor capable of producing a Hadamard transform of the incident optical image has been developed. The device can be operated so as to produce a normal video signal. In a second mode, the device produces four 1 x 4 Hadamard transforms in parallel. Minimal off-chip hardware is required to produce a 4 x 4 two-dimensional transform. This approach offers an opportunity to reduce the size and power requirements of on-board electronics in mini-RPV and guided weapon antijam video data link applications by performing the transform processing function of the airborne encoder directly on the image plane. Performance of the device has been verified by completing the two-dimensional transform of various images, performing the inverse transform, and displaying the resultant images.

TV bandwidth reduction is required to achieve jam resistance in guided weapon and remotely piloted vehicle applications. One way to achieve bandwidth reduction is to eliminate redundant information prior to transmission. The high degree of redundancy in most images can be reduced by employing transform techniques to uncorrelate the scene information. A number of coding techniques can then be used to allocate the available bandwidth to the significant spatial information. A number of transforms, such as the Hadamard, Fourier, Haar, Cosine, and Slant, could be used. Until recently, however, the real-time realizations of these transformations were questionable. The use of large numbers of logic, memory, and linear devices would be necessary to process the video signal in a conventional manner. Charge-injection device technology can reduce these requirements. This technique can be used to combine such widely diverse functions as imaging, memory, and signal processing into a single device to achieve low system cost, size, weight, and power consumption.

The process of computing the Hadamard transform requires linear combinations, sums and differences, of the picture information. This process can be easily mechanized on the focal plane of a CID image sensor by sensing combinations of the signal charge from a number of image-sensing sites (pixels) rather than the individual pixel magnitudes.

Figure 1 illustrates the silicon surface potential and surface charge location under various CID operating conditions. The magnitude of charge stored at each sensing site can be detected by driving one electrode and sensing the displacement current that flows in the other electrode upon charge transfer. The charge can be read in either the positive or negative sense depending upon the direction of charge transfer. This readout operation is nondestructive and can be performed repeatedly with no detectable charge loss. Charge can be removed from any sensing site by driving both electrodes to the minimum voltage magnitude to cause charge injection. This minimum applied voltage is slightly greater than the MOS threshold voltage, to avoid charge loss through interface state recombination (charge pumping). Consequently, bias charge is retained after injection.

Signals proportional to the linear combination of charge from a number of sensing sites can be read out as shown in Figure 2. Prior to readout, the initial condition shown at the top of the figure causes charge at the first two sites to be stored under the column electrodes, and at the third and fourth sites to be stored under the row electrodes. The row line is then set to a reference voltage, allowed to float, and a sample of the floating row voltage is taken to measure this initial condition. If all column voltages are then driven as indicated, charge will simultaneously transfer from column to row at the first two sites, and from row to column at the third and fourth sites. The net change in charge under the row electrodes is then $q_1 + q_2 - q_3 - q_4$, and the change in row voltage is equal to the net change in charge divided by the row capacitance. Any combination can be read by using this technique.
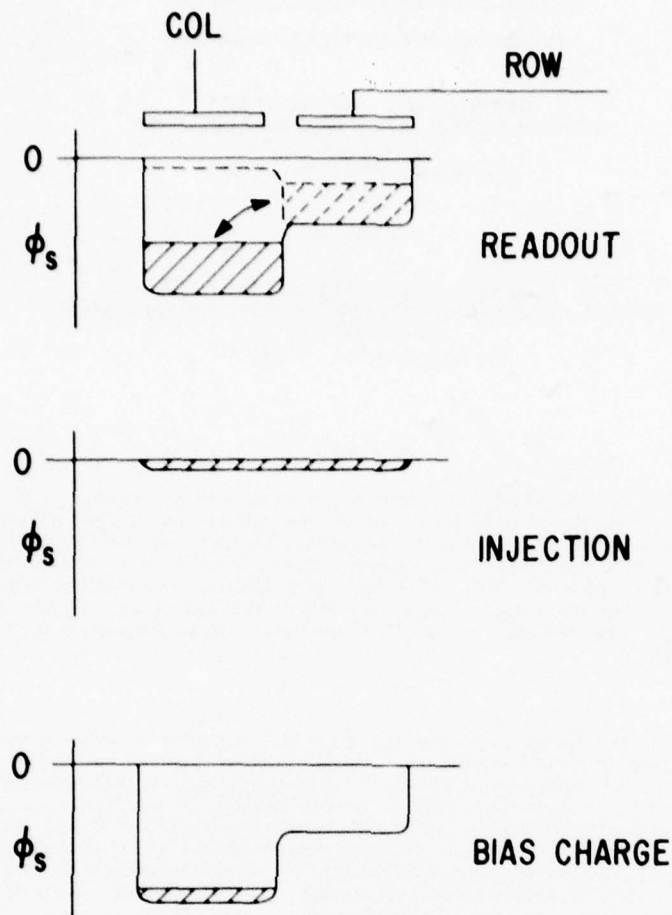
Figure 1. Sensing–Site Surface Potentials
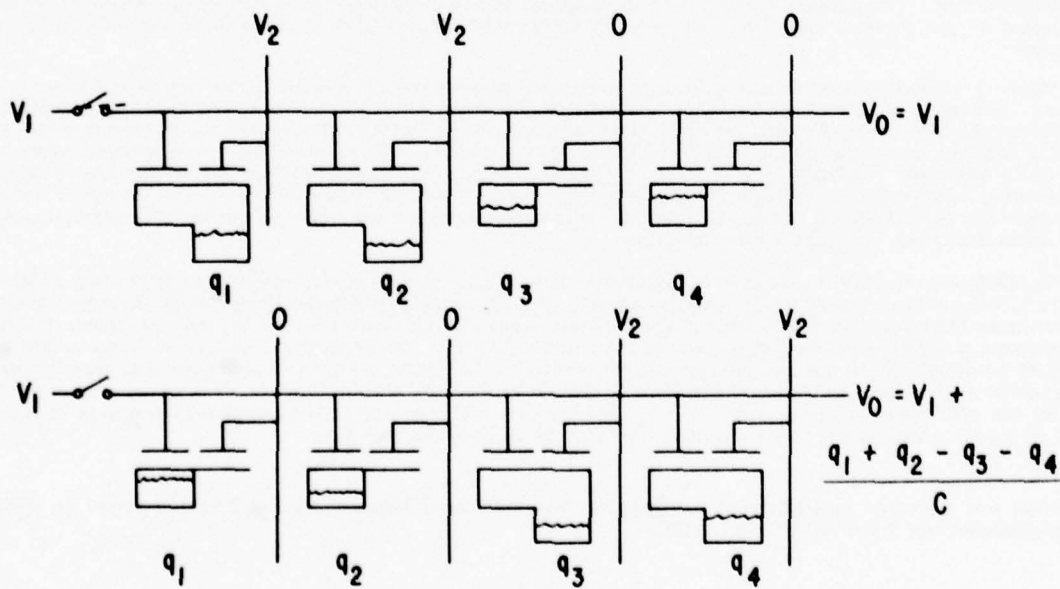
Figure 2. Charge Summation Principle

A 128 x 128 pixel array, designed to perform this signal processing function, is operational. It is capable of producing four 1 x 4 Hadamard transforms in parallel, which are then summed to produce a 4 x 4 transform on one subpicture. The process is repeated to produce a complete image transform. The 4 x 4 format was a design choice for this particular chip; it is not a restrictive requirement.

The approach used to read out 4 x 4 Hadamard transform data is illustrated in Figure 3. The rows selected for readout are connected to four differential amplifiers. Selected array columns are then driven to cause the desired signal summation to occur on all row lines simultaneously. Direct coupling of the column drive pulses to the row lines is cancelled by providing a compensation line, shown at the top of Figure 3, to couple only the column drive interference to the negative inputs of the differential amplifiers.
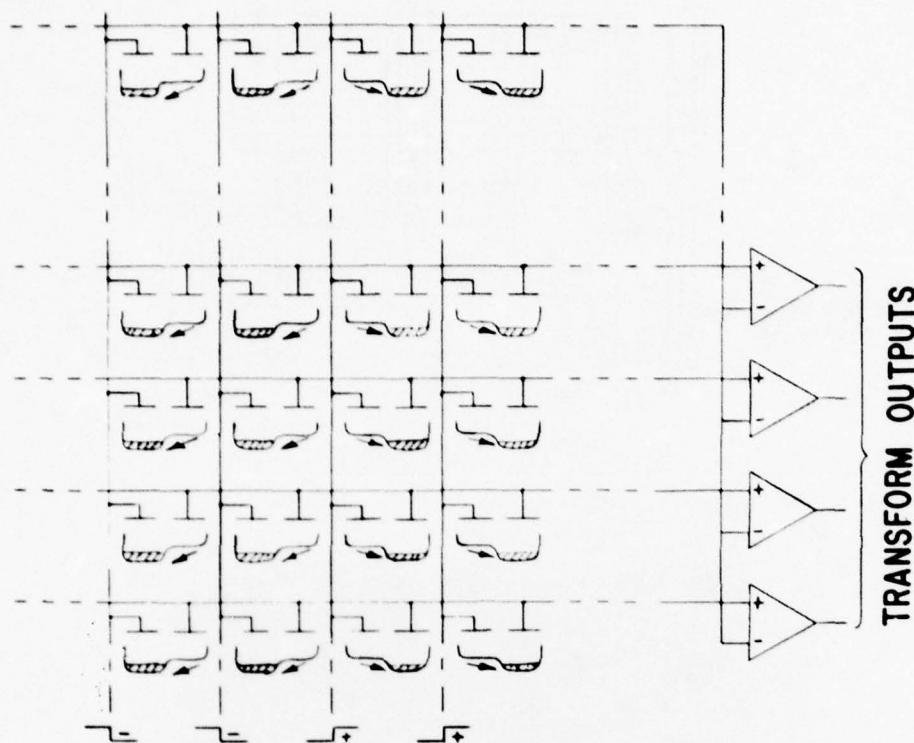
Figure 3. Transform Readout Mechanization

The array organization is shown in Figure 4. Scanning registers are used to select columns and rows in groups of four. Drive signals are applied to inputs E_1 through E_4, and the four selected rows are routed to four on-chip preamplifier transistors. The fifth transistor, output O1, is connected to the compensation line. After readout, the video amplifiers are disconnected from the array by switching the DG input off and driving the selected rows through the ID input, and the selected columns through the E_1 - E_4 inputs, to cause injection.

For transform readout, the drive signals applied to inputs E_1 through E_4 are in the form of a Hadamard code. Alternately, a block code can be applied to these inputs and the outputs suitably multiplexed to effect the more familiar line-sequential video format.

A photomicrograph of the array is shown in Figure 5. The array layout corresponds directly to organization drawing of Figure 4. Sensing site size is 30 x 30 microns, and the overall chip size is 4.82 mm x 5.33 mm.

Operation of the array in the transform mode was verified by digitizing the output samples and completing the 4 x 4 transforms by means of a microprocessor. The inverse transform was then performed, off-line, and stored. Figure 6 shows the reconstructed images for various operating conditions. The normal readout (center) image is that produced by the line-sequential readout operation, while the image directly above center is the result of a complete transform-readout/inverse-transform operation. The remainder of the images of Figure 6 indicate the effect of deleting low-amplitude spatial components. The reproduced images do not indicate the degree of bandwidth compression that can be achieved, since neither proper quantization nor coding was used in this experiment. The analog amplitude of each spatial frequency component of the 4 x 4 Hadamard area transforms was compared with a threshold level and set to

Fig. 9   4 kbit SPS array

4.10-4
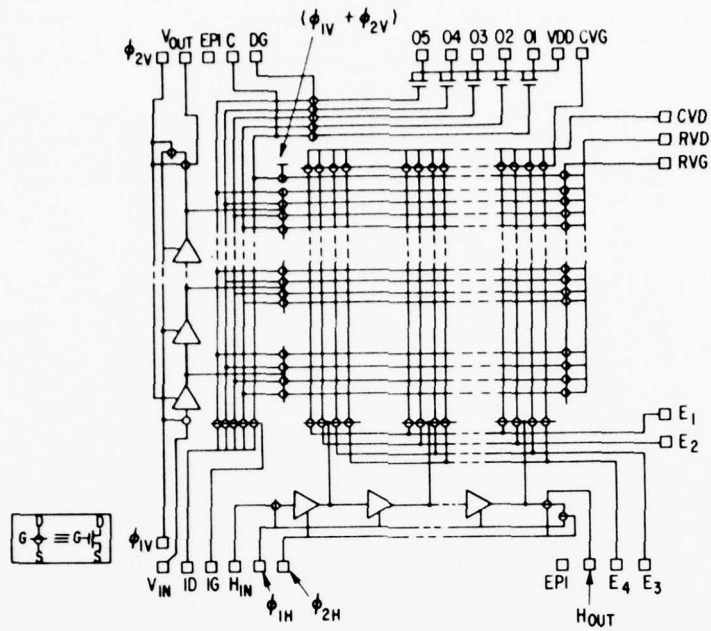
Figure 4. Array Schematic
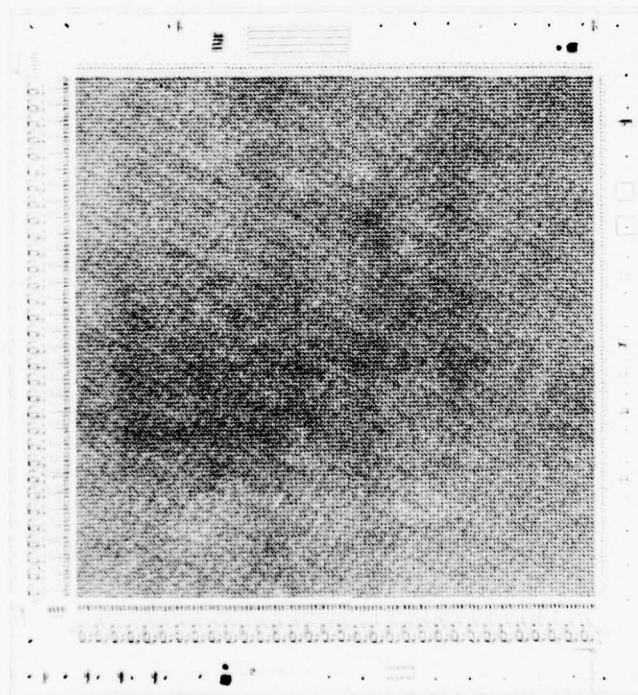
Figure 5. Array Photomicrograph
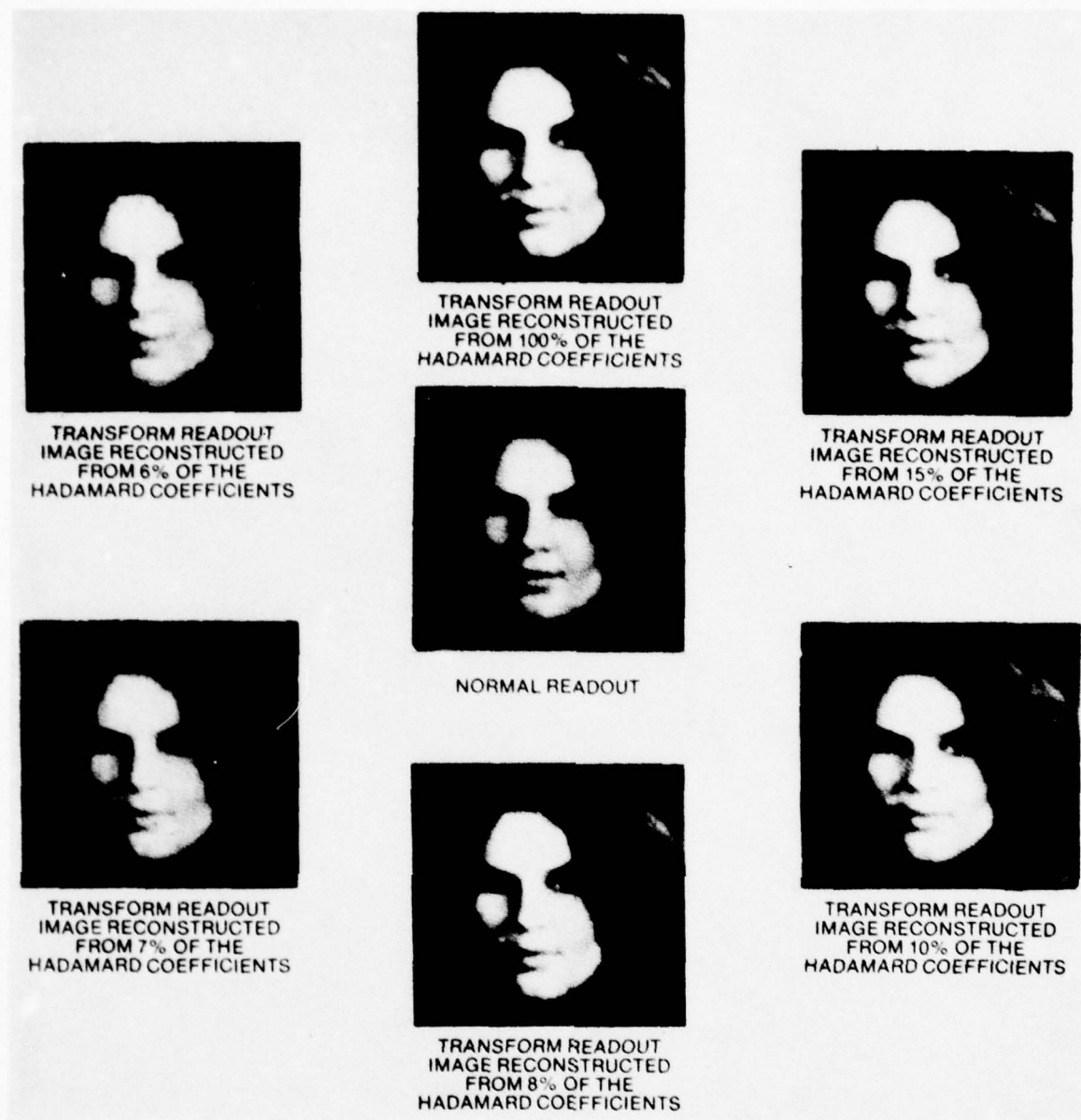
Figure 6. Images

zero if below the threshold. The number on nonzero components used in each reconstruction is indicated below the images, expressed as a percentage of the total number of transform values.

The feasibility of real-time transform coding of an optical image directly on a CID imager has been conclusively demonstrated. This development is considered to be an important first step toward the realization of miniature, low-power, bandwidth-compression (jam-resistant) systems suitable for mini-RPV applications. Other potential applications are in video-guided weapons and cannon-launched guided projectiles (smart bombs).

A HIGH PERFORMANCE CCD LINEAR IMAGING ARRAY

by

A.A. Ibrahim, J.J. White, K.G. McQuhae, K.Y. Yu, D. Gallant
Bell-Northern Research, Ottawa, Canada
and
W.C. Bradley and D.W. Colvin
Itek Corporation, Lexington, Massachusetts, U.S.A.

ABSTRACT

This paper will highlight the development of high performance CCD linear imagers jointly carried out by BNR and Itek. The imager was fabricated with BNR's two-level polysilicon buried channel CCD process. The sensors are MIS buried channel capacitors 8 µm x 18 µm in size with a centre spacing of 12.5 µm. Photocharge collected in the odd and even sensors are read out via two serial shift registers which are clocked synchronously. The novel design technique involved in synchronized clocking systems will be described. Another feature of the design is the ability to vary the integration time independently of the clock frequency. Thus sensor blooming control is easily implemented. Two layers of aluminum are used for both interconnection and light screening of the non-sensor areas. Test results will be described and will include noise spectral density, MTF curves and transfer efficiency measurements. The measurement of noise performance shows considerable improvement in comparison to similar available imagers. Pictures taken using the imagers will be shown.

A high performance linear imager with flexible operating features has many potential applications. These include electronic cameras such as push-broom, rotating panoramic and framing cameras. The system requirements of some of these applications will be discussed.

1. INTRODUCTION

Itek Corp, Lexington, Mass., U.S.A., and Bell-Northern Research, Ltd., Ottawa, Ont., Canada have been working since 1976 on the development of a family of CCD imaging devices intended for application in military reconnaissance systems. We report here on the first of this family, a 256 element CCD linear imaging array. With respect to available commercial devices, this chip, designated 631, has several attractive features. It has some differences in architecture for convenience, it has better technical performance in some key areas such as noise level and dynamic range, and is characterized under conditions representative of military, rather than commercial, application.

This paper will discuss three topics. We will describe the architecture and fabrication of the chip, and present test results including pictures covering the radiometric dynamic range. Finally, we will briefly discuss the application of this chip to military reconnaissance systems.

2. DEVICE ARCHITECTURE AND FABRICATION

A schematic of the linear imager is shown in Fig. 1. The linear imager consists of three building blocks; a) a row of 256 MIS sensor elements (pixels), which are defined by a serpentine field oxide channel stop and a polysilicon photogate, b) a CCD shift register located on each side of the sensor row to transport the photo-charge to the output circuit and c) a sloshing (transfer) gate on each side of the sensor row to transfer the charge from the sensor to the CCD. This bilinear arrangement of MIS sensor/CCD shift registers allows close packing of the sensor elements. In our linear imager, the active sensor area is 8 µm x 18 µm and the sensor to sensor spacing is 12.5 µm, which gives a spatial resolution of 40 line pairs/mm.

During operation of the imager, a voltage is applied on the photogate, creating potential wells in the sensor sites. Photo-generated electrons are collected in these potential wells. After a suitable integration time, the collected signal charge is transfered to the CCD shift registers via the sloshing gates. The proper timing clocks are applied on the sloshing gate (from low to high state) and on the photogate (from high to low state) to allow the signal charge to transfer in a short time of the order of 20-50 nsec. A schematic of the electrostatic potential profile under the photogate, sloshing gate and CCD electrodes during charge integration and transfer is shown in Figure 2. The serpentine channel stop directs the charge from the odd number sensor elements to the upper shift register and the charge from the even number sensors to the lower shift register (Figure 2). Since both sloshing gates are pulsed on at the same time it is necessary to have CCD electrode of the same phase facing the odd sensors on the upper side and the even sensors on the lower side. This is readily implemented by offsetting the two shift registers by half a CCD stage of delay, as shown in Figure 1. After the charge is transferred from the sensor element into the CCD storage electrode, the CCD shift register could be clocked in the normal two phase or four phase fashion to shift the charge packets to the two output circuits, where the data are read out serially. To maintain the proper sequence of the data during readout, the lower shift register is half a stage shorter than the upper shift register (127 1/2 stages compared with 128 stages), which enables signal charge to be read out of the upper shift register during $\emptyset_1$ timing and the lower shift register during $\emptyset_2$ timing. Now the output data is readily multiplexed off chip to obtain the complete stream of 256 data information. The timing clocks and output waveform are shown in Fig. 3. It is obvious from the previous discussion that the simple arrangement of offsetting the two shift registers and making one of them 1/2 stage shorter simplifies the clocking immensely. This is in fact one of the attractive features of the 631/BCCD2 linear imager.

The output sensing circuitry of the shift registers is shown in Figure 4. It consists of a presettable floating diffusion connected to an on chip pre-amplifier (depletion type MOST source follower). The photocharge is sensed as a voltage change at the output of the source follower.

The imagers were fabricated using BNR's two levels polysilicon gate buried channel technology. The process allows the on-chip incorporation of enhancement and depletion (ion implanted) MOST structures. This attractive feature contributed to the low noise performance of the output pre-amplifier. The process uses three ion implants steps plus n+ diffusion for source and drain. It also features isoplanar oxide isolation and two levels of aluminum interconnects.

3. TEST PROCEDURE AND EXPERIMENTAL RESULTS

3.1 GENERAL INFORMATION

The devices were evaluated on two levels; normal device characterization tests such as transfer efficiency, spectral response, preamplifier gain and noise measurements performed at BNR, and extensive testing at Itek to characterize the device as an image transducer in a situation comparable to a high quality reconnaissance system application. The data produced provides the system designer with component data in order to confidently estimate the performance of a hypothetical system and provides feedback data to the CCD fabrication laboratory in order to improve future designs.

High quality airborne military electronic reconnaissance camera systems are rather complicated, with the data stream from the CCD or other solid state array typically undergoing sophisticated signal processing in the aircraft, digital transmission or recording, and digital data processing on the ground before display. In order to fairly characterize a CCD array for this application, we must duplicate most of those functions with test gear such that the final displayed output is limited in its quality by the device under test, rather than by the test equipment. In the case of imaging CCDs, this is a difficult problem due to the low noise and high dynamic range.

3.2 TEST PROCEDURE

Figure 5 is a block diagram of the test set used for the characterization testing of high performance image transducers. It shows the image producing process from the input scene to the final output, which may be a hard copy print of the image seen by the CCD or a computer printout representing the analysis of some properties of the image. At the left side, we show the various input possibilities for a CCD under test. We may choose among real scenes, model scenes, photographic reproductions of real scenes, or standard test charts. This test input is imaged on the CCD under test with a lens whose imaging performance is known. The image transducer in our situation is a CCD under test with its operating electronics. In order to simulate high quality data recording, we use a Data General Eclipse S-200 mini-computer with a customized set of peripheral equipment. The pixel stream from the CCD is digitized by a high quality A/D converter capable of 10^7 conversions per second with 10 bit resolution and fed into a semiconductor memory which can store up to 7×10^6 10 bit words. This is sufficient to reproduce frames of imagery large enough for display and characterization. The S-200 can then retrieve the data from the semiconductor memory, perform various manipulations of the data, and then write it onto tape for more detailed analysis in an IBM 370/150 computer which is a general purpose off line machine. That computer can perform data manipulations of several types. It can perform image enhancement (such as calibrating out the element to element variations that exist in virtually all arrays) for the purpose of presenting hard copy image display. It can perform a number of statistical analysis of the imagery such as signal to noise ratio calculations. It can also alter the recorded data to simulate the effect of other system components so that complete systems may be simulated for system design purposes. Finally, we have available as output peripherals, a choice among a laser writer (for hard copy imagery), a line printer for tabular data, a plotter for graphical data, or a tape deck for tapes to be stored or transmitted to other organizations.

3.3 EXPERIMENTAL RESULTS

Table I is a summary of the properties of the 631/BCCD2 linear imager. The transfer efficiency is of the order of 0.99999 up to 10 MHz CCD clock rate. Transfer efficiency of this value allows an array of this architecture to be built with 3000 elements or more before the transfer efficiency cause a noticeable loss in resolution. The radiant sensitivity values quoted are typical of front side illuminated CCDs where the incident light must pass through one poly silicon gate before entering the silicon substrate to create photocarriers. The peak quantum efficiency is about 70%. The noise level given is the rms of the additive random component which is superimposed on the recorded image. In addition, there are systematic noise sources such as elemental variations in dark current level (3% rms) and optical sensitivity (3% rms for all wavelengths between 0.4 to 0.9 μm). These systematic noises are very low and will not affect the imager performance under normal operating conditions. Nevertheless, to obtain ultimate imager performance they are removed by system signal processing. Computer tallies of average and rms are then made for rows and columns in the image. The correlated double sampling effectively removes the noise associated with the resetting of the floating diffusion capacitance at the output and also attenuates the 1/f component of preamplifier noise. Combined with the sensitivity to sunlight (used for calculations since most applications involve sunlit scenes) the noise level amounts to an exposure level of 0.25×10^6 j/m^2, which is lower by a factor of several than is achievable with other available devices.

The saturation level, defined as the signal level where image quality begins to degrade seriously, is 500,000 electrons per pixel, giving a dynamic range of over 10,000. The real importance of this noise level is that for signal levels of 5,000 electrons per pixel or higher, the imagery is limited only by the statistical of "photoelectron" noise in the detected photocarriers from the scene. The imaging array does not degrade signal to noise. When a typical system performance budget is established, including the effects of low contrast in a normal aerial scene and system Modulation Transfer Function (MTF), a typical signal level required to produce useable imagery is 5,000 electrons per pixel. The noise data are for an operating rate of 10^6 pixels per second.

Above, we mentioned that the preamplifier noise has a 1/f component in it. Figure 6 shows a measured noise spectrum. This spectrum is measured with the CCD clocks turned off so that sources of noise associated with the signal current, dark current, and clocking are not present. The noise remaining is that associated with the output transistor. The spectrum shows a 1/f noise corner at a few hundred KHz and a preamplifier rolloff near ten MHz. The use of correlated double sampling effectively suppress the 1/f component at the expense of increasing the flat band noise voltage component by $\sqrt{2}$. For low pass filtering at about 1.5 MHz (appropriate to the operating rate) an integration of the spectrum yields a noise level of 30 electrons per pixel rms. This gives reasonable correlation with the computer calculated noise at low exposure, where preamplifier noise should dominate. Also shown in Figure 6 for comparison, are noise spectra curves measured on preamplifiers incorporated into CCDs, from other manufacturers. These measurements were carried out over the last two years by ITEK corperation using the same measurement set-up.

Signal and noise transfer curves are shown in Figure 7. They show the average response of the CCD to changes in exposure ranging from the noise level, which sets an effective limit in both the ability to produce imagery and on the ability to measure the average signal, to the saturation level, where image quality begins to degrade due to smearing in the CCD shift register. The rms noise is limited primarily by preamplifier noise at very low exposures, but by photoelectron noise at most exposure levels of practical interest, as was noted above. A chip gain value of about 1 µv per electron is appropriate, as may be verified by noting that with this gain constant, the photoelectron noise is the square root of the signal level as predicted from the Poisson statistics of the detected photoelectrons.

The best test of an imaging system is imagery. Figure 8 shows a sequence of pictures taken with the 631/BCCD2 device using the computer controlled imaging equipment described earlier. These pictures have been processed to remove elemental variations in sensitivity and dark current, but were taken with conventional (single sample per pixel) rather than with correlated double sampling. From signal levels near saturation down to less than 10^{-2} of saturation (5,000 e$^-$ signal) very little difference may be discerned. At 10^{-3} of saturation, (500 e$^-$ signal) noise begins to be visible to the eye. At 1×10^{-4} of saturation (50 e$^-$ signal) the picture is noticeably noisy and represents about the threshold of usefulness. The test target was a high contrast target in all cases. At the lowest levels, we also notice some loss of definition, apparently showing the ubiquitous loss of transfer efficiency in CCDs when transferring very small charge packets of a few hundred electrons or less. The streaky nature of the noise in the pictures dominated by pre-amplifier noise is due to the 1/f spectrum of that noise.

4. APPLICATION TO MILITARY RECONNAISSANCE

The area of application that we are addressing here is the recording of high quality aerial imagery from high performance aircraft. There are two prime modes of operation that are usually considered for the newly emerging CCD array technology. They are usually referred to as "push broom" operation or "pan" operation. Figure 9 shows the operating situation for push broom operation. Here, a camera remains in a fixed position relative to the aircraft while the aircraft moves to provide the scanning. The format of the array, shown projected onto the ground, moves along the ground like a pushbroom, sweeping over the scene to be imaged. The other concept places the array in the focal plane of a panoramic camera. This camera structure rotates with respect to the aircraft and sweeps out a strip of imagery which is perpendicular to the flight path (Fig. 10). The pan camera covers more area on the ground, but, because of its much higher scan rate, has much less exposure time. Table II illustrates a sample signal to noise calculation for a typical situation. Here, a high performance aircraft is assumed to be flying over an area of interest. A camera with f/4 optics and 0.4 m focal length has a linear CCD imager in the focal plane. Signal to noise calculations for the two camera configurations have been made using a typical set of conditions for illustrative purposes. The signal to noise ratio is defined in relation to the imaging of a high contrast black and white bar pattern. The signal is the mean difference in the signal between the blacks and whites (peak to peak). The noise is the rms of the random fluctuations due to noise. To be barely visible, the USAF 1951 Resolving Power Test Target requires a signal to noise ratio so defined of about 2 to be clearly resolved. The calculations indicate that for 0.35 m sample spacing on the ground, a push broom camera has sufficient exposure with some margin to produce good imager. The rotating pan camera was assumed to have a 6 degree field of view and a scan efficiency such that while recording the angular scan rate was 1.5 radians per second. The estimate of signal to noise ratio is such that this camera would provide imagery under these conditions which is marginal at best. The main reason is that the exposure time in the pan camera is smaller by a factor of about 30. Thus, we see by illustration that the linear CCD has the capability of being usefully designed into some, but not all, airborne camera applications. For those applications where a linear imager does not have enough sensitivity, other CCD configurations, particularly areas CCDs designed to operated in the Time Delay and Integrate mode are needed for useful performance.

5. SUMMARY

To summarize, close co-operation between Itek and BNR results in the fabrication and extensive testing of very high performance linear imagers. These devices were shown to be capable of taking pictures over the entire radiometric dynamic range of 80 dB. An application analysis shows that the imaging device will give good imagery with a pushbroom aerial reconnaissance camera system.

BNR 631 B-CCD LINEAR IMAGER BLOCK DIAGRAM
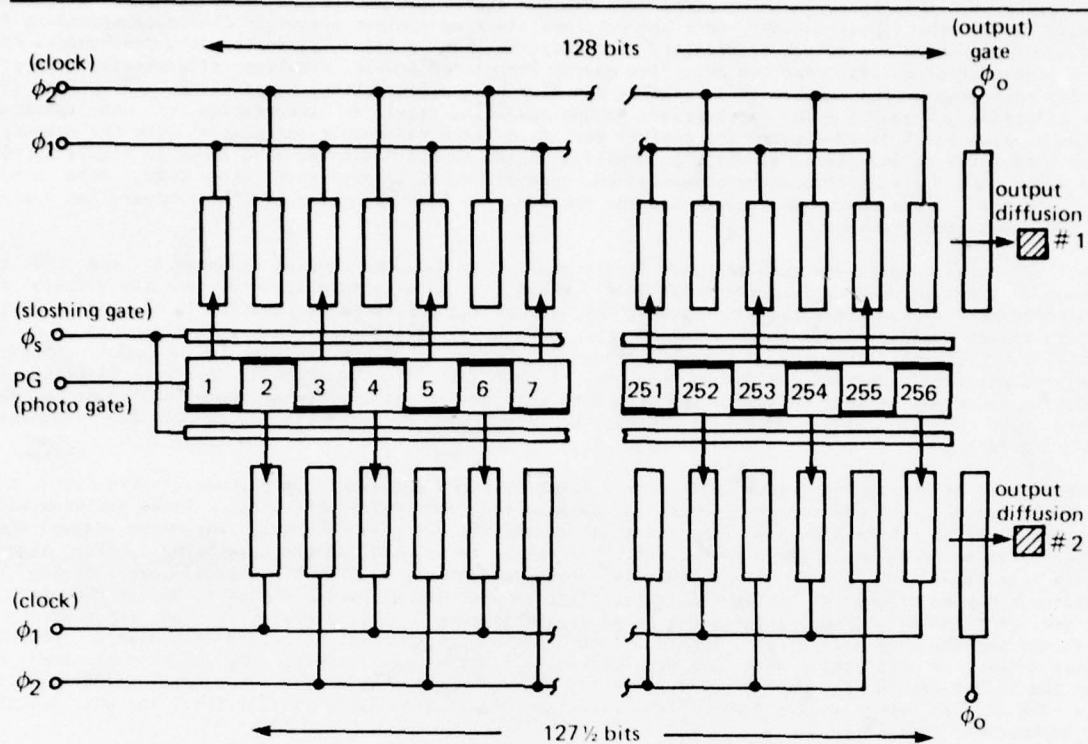
Fig. 1   Schematic of 631/BCCD2 Linear Imager.

CROSS SECTION OF IMAGER AND POTENTIAL PROFILE
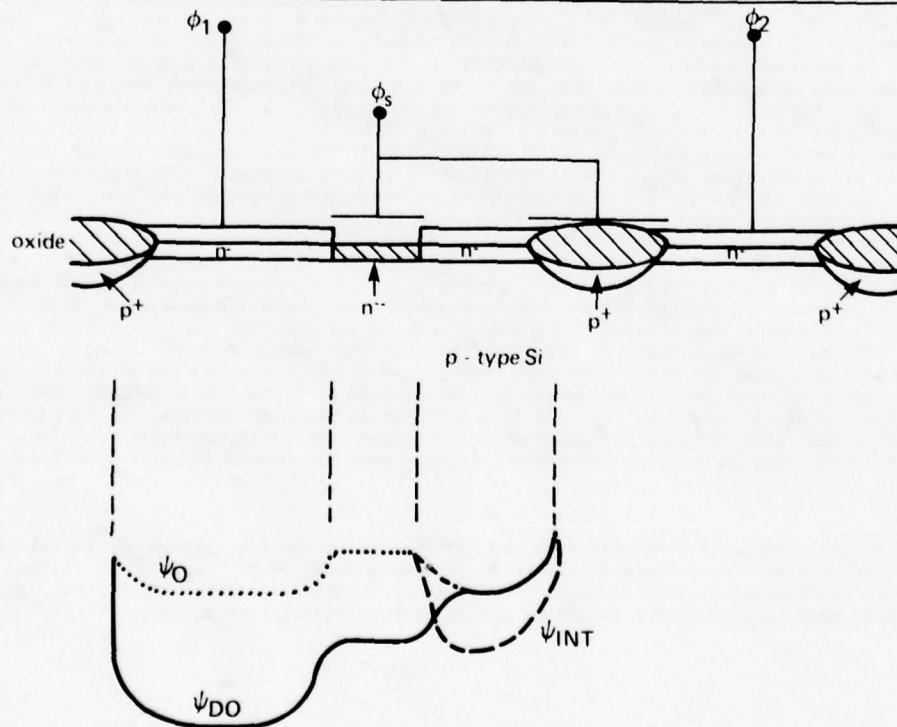
Fig. 2   Potential profile under the photogate, sloshing gate and CCD gate during charge integration and transfer.

BNR 631 LINEAR IMAGER TIMING DIAGRAM
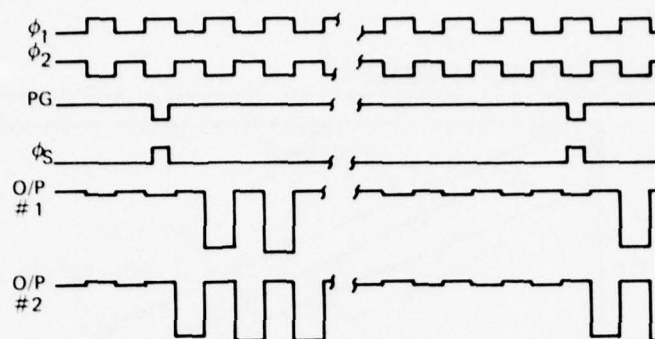
Fig. 3   Clock and output waveform of linear imager.

OUTPUT CHARGE SENSING CIRCUIT
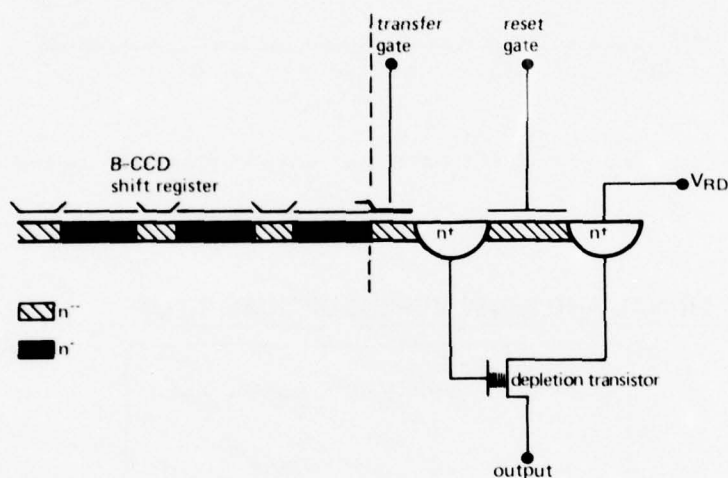
Fig. 4   Schematic of the output circuitry of the shift register.

PRODUCTION OF REAL OR SIMULTATED E-O IMAGERY
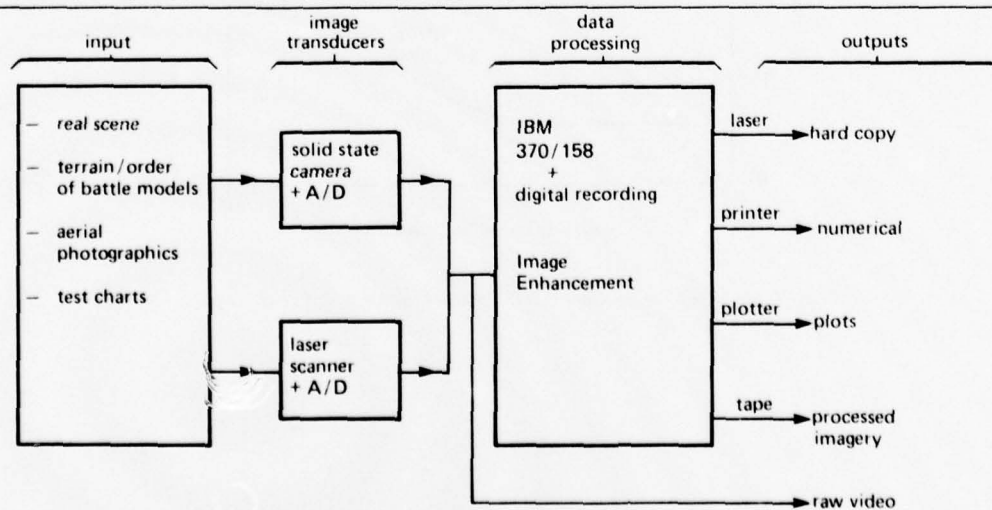
Fig. 5   Block diagram of the imager testing system.

NOISE SPECTRAL DENSITIES
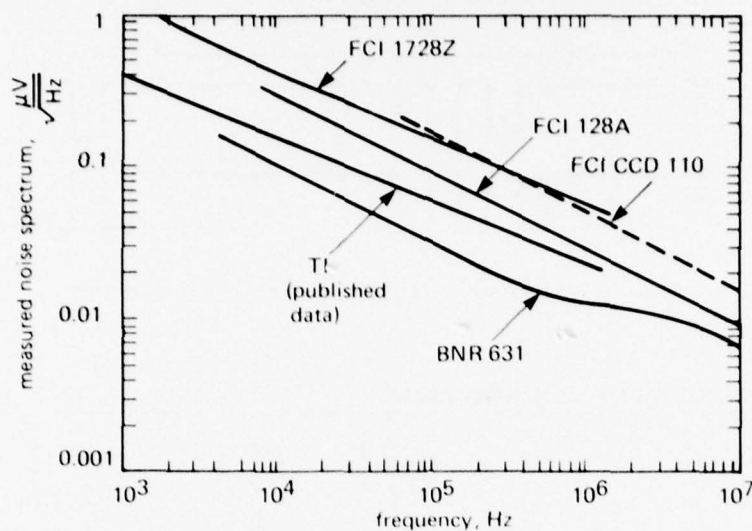
Fig. 6   Preamplifier noise of the 631/BCCD2 linear imager compared with available devices.

SIGNAL AND NOISE VERSUS OPTICAL INPUT
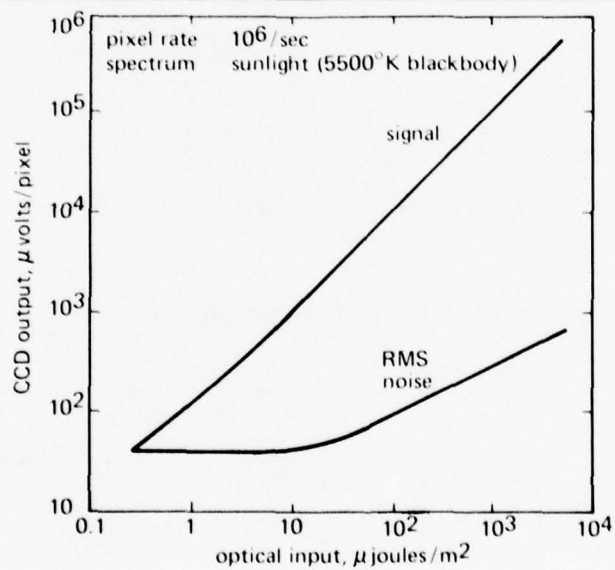
Fig. 7   Transfer curve of CCD output versus optical input.

SAMPLE IMAGERY FROM 631/BCCD2 CHIP
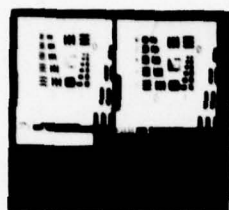
Near
saturation
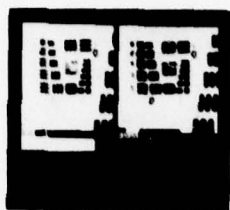
0.1
saturation
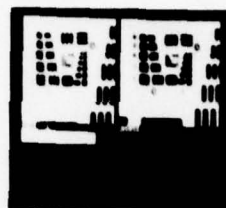
0.01
saturation
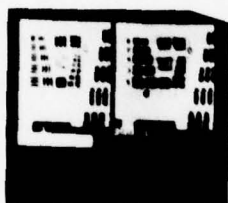
0.001
saturation

0.0001
saturation

Fig. 8 Sample imagery from 631/BCCD2 device.

PUSHBROOM SCAN FOR ONE-DIMENSIONAL ARRAY
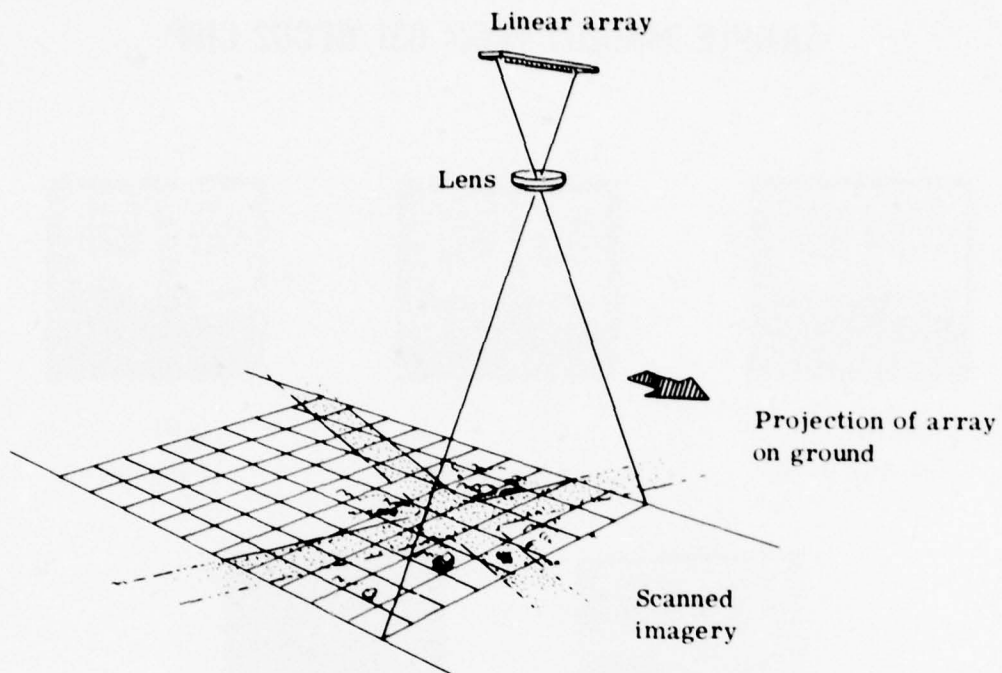
Fig. 9   Pushbroom camera in operation.

PANORAMIC SCAN FOR ONE-DIMENSIONAL ARRAY
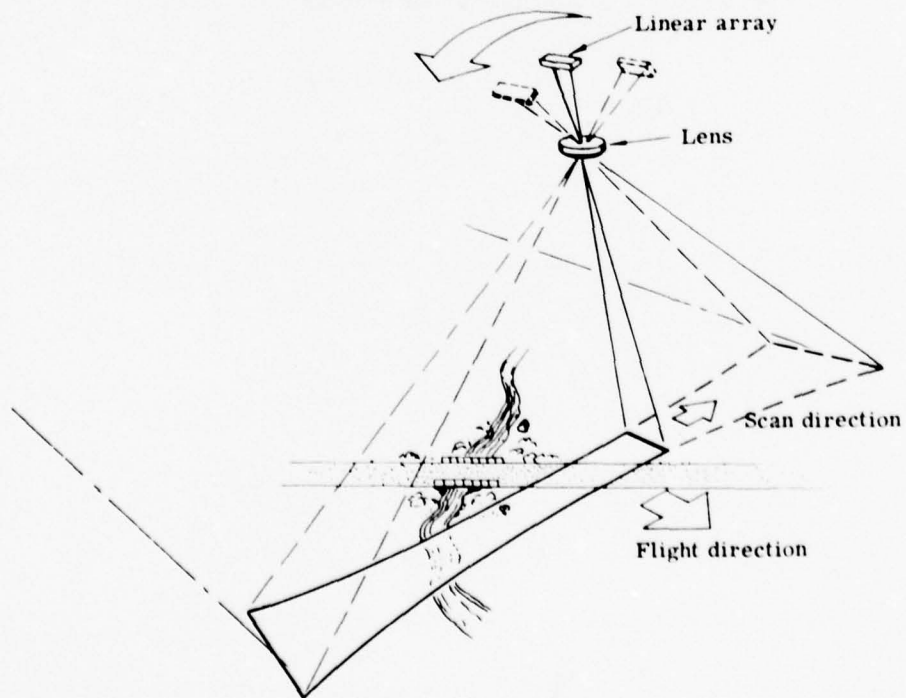
Fig. 10   Rotating panoramic camera in operation.

SUMMARY OF PERFORMANCE DATA

Number of elements	256
Element spacing	12.5 μm
Electrode configuration	4 ϕ (can be clocked as 2ϕ)
Transfer efficiency; optical imager	0.99995
Sensitivity electrical input	0.99999
Peak (0.75 μm)	0.4 A/W
Sunlight (5500° K blackbody)	0.1 A/W
Noise level (with correlated sampling)	40 e⁻/pixel
Saturation level	500,000 e⁻/pixel
Noise equivalent exposure	2.6×10^{-7} J/m^2
MTF at Nyquist limit (40 1p/mm)	
Green light	0.55
Sunlight (5500° K blackbody)	0.25
Dark current	1 to 3 nA/cm^2
Dark current non uniformaties	$<$ 3% RMS

TABLE I: Summary of Linear Imager Performance Data.

IMAGING PERFORMANCE ANALYSIS

Definition: SNR = $\dfrac{\text{Peak-to-peak signal from black/white bar chart}}{\text{RMS signal from uniformly illuminated pixels}}$

Conditions:

Illumination	100 W/m^2 (low or hazy sun)
Reflectivity	0.2
Scene contrast	0.2
Transmission (atmosphere and optics)	0.4
Optics speed	f/4
Optics focal length	0.4 m
System modulation transfer function	0.2
Aircraft velocity	600 m/sec
Aircraft altitude	12,000 m
Sample spacing on ground	0.35 m
Sample spacing on array	12.5 μm

Analysis results for bar chart at Nyquist Frequency
(one cycle per two pixels)

Pushbroom camera (integration time = 0.5 msec)	SNR = 3.7
Rotating panoramic camera (integration time = 16 μsec)	SNR = 0.26

Conclusions:

- Pushbroom camera gives good imagery with linear array

- Rotating panoramic camera gives marginal imagery with linear array; a TDI or other configuration is needed

TABLE II: Imaging Performance Analysis

DISCUSSION

J.J.Stapleton

1 — Isn't 0.4 A/W responsitivity close to 100% quantum efficiency?

2 — At what wavelength?

3 — What is the input capacitance of the preamplifier transistor?

Author's Reply

1 — No, about 70%

2 — 0.75 microns

3 — 0.25 pF.

Schaeffer

Could you tell us your idea concerning the practical production of your device?

Author's Reply

Devices not too different from this are available commercially. Their performance, particularly in term of noise, is sometimes deficient. There are no plans for high production.

DEVELOPMENT AND APPLICATION OF AN S.A.W. CHIRP-Z TRANSFORMER

M.B.N. Butler, B.Sc., PhD., M. Inst. P.,

Microwave and Electronic Systems Limited,
Lochend Industrial Estate,
Newbridge,
Midlothian,
EH28 8LP,
Scotland.

1. INTRODUCTION.

With the introduction of increasingly sophisticated signal processing techniques methods of obtaining Fourier Transforms or Spectrum Analyses quickly and accurately have attracted much attention. The bipolar digital F.F.T. has made very rapid progress both in chip integration of multipliers and butterfly functions and in speed. However, the present state of the technology suggests that even spectral analysis of a range containing one hundred elements to six-bit accuracy becomes very power consuming if the process is to be completed in 50µs.

In this paper a development is reported in which the Chirp-Z transform algorithm is truncated to perform spectral amplitude analysis. The basic process is described in relation to a unit giving in excess of one hundred resolved spectral lines with a resolution of 50kHz and a total processing time of 50µs per data set. The displayed spectrum covers a dynamic range in excess of 36dB (six bits) essentially linearly and without process induced spurious signals.

Two distinct applications of the unit are considered; to a coherent microwave radar in which Doppler spectral analysis is employed and to signal analysis in a 10µ wavelength laser wind velocity equipment. In both applications the high speed of the unit has proved to be particularly significant.

2. THE CHIRP TRANSFORM.

The chirp transform is a signal processing algorithm which can perform Fourier Transformation in real time. It relies on structuring the conventional Fourier Transform as two complex multiplications separated by a convolution stage as shown in figure 1. The input signal $r(t)$ is first multiplied by a linear chirp (C_1) which may be represented by:

$$C_1 = \exp(j\mu t^2/2) \qquad \qquad ....1$$

the resultant signal is convolved with a second linear chirp (C_2) which is assumed to have frequency dispersion which is matched to the sideband selected after the first multiplication. After this convolution the signal is:

$$R(\mu\tau) = \int r(t) \exp(j\mu t^2/2) \exp(j\mu(t-\tau)^2/2). \, dt.$$

$$= \exp j\mu T^2/2 \int r(t) \exp -j(\mu\tau)t dt. \qquad \qquad ....2$$

A further multiplication by $C_3 = \exp -j\mu\tau^2/2$ gives:

$$R^1(\mu\tau) = \int r(t) \exp -j(\mu\tau)t \, dt. \qquad \qquad ....3$$

from which it can be seen that the output $R^1(\mu\tau)$ is the Fourier Transform of the input function with a transform variable $\omega = \mu\tau$.

In essence this process relies on the fact that in a linear chirp delay time and frequency within a discrete band are related by the dispersive slope μ.

In practice all chirps will be of finite duration and it can be shown that the bandwidth $B_2 = \mu T_2$ must be equal to the sum of the input bandwidth B_s and the bandwidth of the premultiplier, i.e.

$$B_2 = \mu T_1 + B_s = \mu T_2 \qquad \qquad ....4$$

Since the convolution process will only occur for a time period T_1 the output function for each inband frequency component of the input signal will take the form $\text{sinc}\{(\mu\tau) T_1\}$ which has a 4dB width of $1/T_1$ and also introduces sidelobes 13dB below the main peak. These sidelobes can be reduced at the expense of frequency resolution by spectral weighting as in linear chirp pulse compression (Cook and Bernfield(1967)). Correct sidelobe suppression can be implemented by means of spectral weighting after the post-multiplier.

In circumstances where the post multiplier is omitted as in figure 2 the amplitude spectrum of the input signal can be obtained by amplitude detection of the output from the convolver (C_2). At this point the amplitude spectrum is correctly obtained but the function $R(\mu\tau)$, equation 2, carries a multiplicative term $\exp(j\mu\tau^2/2)$ which indicates that each frequency component is on a carrier which varies in direct proportion to the time of occurance. This arrangement does not allow the use of spectral weighting since the post multiplier is omitted. An alternative is to use spectral weighting in the pre-analysis stage and this has been employed in the system now considered.

3. SIDELOBE SUPPRESSION.

Theoretically correct sidelobe suppression can be achieved by spectral weighting of the premultiplier chirp (C_1) employing some suitable function such as the Taylor, Hamming or Gaussian functions. These functions are all essentially similar in that they have a smooth roll-off on either side of a centre frequency peak. For a linear f.m. chirp of moderate or large time-bandwidth product the amplitude weighting functions are essentially the same in both frequency and time domains. In the unit now being considered C_1 has T.B. = 100 and the equivalence of the weighting functions will be shown to be good enough to achieve 38dB sidelobe suppression without pulse broadening in excess of that defined by the Hamming function.

The method employed for sidelobe suppression takes account of a number of practical problems,

a) Mixers suitable for I.F. multiplication do not allow simultaneous variation in both input signals if an output linearly proportional to two input functions is required;

b) Analogue multipliers which do achieve a linearly proportional product have limited bandwidth capability and operate with greatly reduced dynamic range for frequencies above 10MHz.

Figure 2 shows a circuit in which the input burst of data, 25µs long is amplitude modulated by a cosine function on a pedestal generated by bias adjustment on the analogue multiplier. The cosine function which is time synchronous with the input data is obtained by reading an appropriate function out of a PROM. In this way the relative magnitudes of the cosine and pedestal amplitudes allow functions similar to the Hamming function to be obtained. Fine adjustments to the function when observing the output produced by a single input tone allows optimum sidelobe suppression and resolution. From the analogue multiplier the signal passes to mixer (M1) which is driven by an upconversion tone at 54MHz or 57.3MHz depending on whether the available bandwidth is to be employed for symmetrical double sideband or single sideband analysis. After upconversion the signal is mixed in M2 with the premultiplier chirp (C_1) which is time synchronised with the data burst. At this point the desired product of amplitude function and premultiplier chirp is achieved.

4. SPECTRUM ANALYSER OPERATION.

The two S.A.W. devices have linear f.m. characteristics in which frequency decreases with time with dispersion of 6.25µs/MHz. The premultiplier is centred on 22MHz and covers 4MHz in 25µs while the compressor is centred on 32MHz and covers 8MHz in 50µs. Both devices employ the λ/6 interdigital configuration which was chosen to minimise finger reflections without introducing the narrow electrodes required in the split-finger configuration. The devices also have cosine-squared time extensions to reduce spectral ripple and the associated spurii in the output spectrum.

Spectral usage is illustrated in figure 3. The choice of an initial upconversion to a band near 54MHz allows the S.A.W. device bands to be chosen with modest fractional bandwidths and to avoid overlapping fundamental and lower order harmonics. This choice is of particular importance in the reduction of spurious signals without having to employ additional filters. As shown in figure 3 the difference frequency at the output of the premultiplier stage is selected by the bandpass characteristic of the compressor. Since the difference frequency band contains components which are spectrally inverted the requirement is for both chirps to be of the same sense.

4.1. Single sideband operation.

The S.A.W. devices employed give a central region of 4MHz in which all spectral components are processed equally. Outside this range the signal output and sidelobe levels degrade progressively as the associated signals after the premultiplier fall partly outside the bandwidth of the compressor. Figure 4 shows how the amplitude response and sidelobe suppression behave over extended bandwidth. Over the 6.5MHz bandwidth offered by employing an upconverter at 57.3MHz the amplitude variation is less than 1dB and the sidelobe degradation is only significant at the edges of the band.

To illustrate the analyser performance figure 5 shows the output with inputs at 4.0MHz and 4.3MHz with 20 dB differential in amplitude. Resolution of 50kHz to -3dB and minimal sidelobes, to -40dB in this range are demonstrated. Figure 6 shows the spectrum obtained from a composite input consisting of a square wave at 300kHz and a 4.0MHz continuous tone suppressed by 20dB. The linearity of response and absence of spurious signals over the full bandwidth of the unit are illustrated.

4.2. Double sideband operation.

To permit optimum use of the analyser in processing returns from a coherent radar the circuit in figure 7 was developed. In-phase (1) and quadrature (Q) returns are amplitude modulated in matched input circuits. The resultant signals are then upconverted with 54MHz tones in quadrature before entering a resistive summing network. Following the summing circuit the output from a single return will lie in the upper or lower sideband range depending on whether the Q return had a phase advance or delay with respect to the I return. In practice if the I and Q returns relate to Doppler signals processing can be arranged to give receding targets on the left of the display and advancing targets to the right.

5. APPLICATIONS.

Many applications have been suggested for a fast, linear spectrum analyser. In this respect the elapsed time from input or display trigger is a very linear function of the input frequency. This linearity relates directly to the chirp linearity required to achieve 38 to 40dB sidelobe suppression and suggests that spectral determination to better than ±0.01% should be possible over the central 4MHz but will require timing measurement to ±2.5ns. Two particular applications are now considered in more detail.

5.1. Coherent pulse Doppler spectral analysis.

Roberts et. al. (1976) and Jacobs et. al. (1976) have previously described the principles of this application. As shown in figure 8 the outputs from a coherent radar after M.T.I. detection are clocked into a C.C.D. or Digital bipolar tapped shift register. In this record each step corresponds to a range cell and similar records are made in the I and Q sections of the processing circuitry. When the first set of data is complete it is shifted into a set of parallel shift registers and the process repeated on successive returns until the complete array is full. Since the data rate in the returns stored in a particular shift register corresponds to Doppler shifts at microwave frequencies, typically less than 20 kHz, a fast readout can be employed from each register to multiply the data into the band of the spectrum analyser. This process allows the serial analysis of all the data in a group of range cells without loss of radar data.

In a particular application to a coherent radar operating at 0.107 meter wavelength the conversion rates are shown in figure 8. A batch of detected data is clocked into the tapped register at 6MHz, i.e. 25 meters between range cells. This data is then passed into the set of parallel stores at the 1kHz repetition rate of the radar transmission. Finally store readout is made at 5.12MHz to give a speed-up factor of 5120 with a total spectral coverage extending over ±391 Hz. With this speed-up factor the S.A.W. spectrum analyser gives a display rate of 31.25Hz/μs and a resolution of 10Hz to the -3dB points on a single tone input. When the display is at 2μs/div. as in figure 9 the effective Doppler shift indicates a target velocity of 3.34m/sec. per division. Figure 9 shows 16 range cells at 6km from the radar which was set with an elevation of 1.3°. The returns are from airborne matter at an altitude of approximately 450ft and demonstrate the phenomenon of windshear. Between range cells 10 and 12 the Doppler frequency change indicates an increase in wind velocity from 2.5m/s to 5m/s.

5.2. Laser wind velocity measurement.

The use of CO_2 lasers at 10.6μm wavelength has been reported by a number of groups (Alexander et. al. (1977), Hughes et. al. (1972 and 1973), Sonnenschien et. al. (1971)) in application to remote wind velocity measurement. The principles of such a system are essentially similar to a homodyne radar or communications system as shown in figure 10. The particular features of this system have been described by Alexander et. al. (1977) and will not be considered in greater detail here. Alexander et. al. (1977) show that the main return occurs from scattering in the laser "focus" shown in figure 11 (a). Such returns are primarily from airborne particulate matter and thus show a combination of Doppler shifts relating to mean wind velocity and the random particulate motion. Figure 11(b) shows an experimental detector output when the particle density and signal return were relatively high. This signal indicates pseudo-coherence over a period of a few tens of microseconds and shows that in favorable conditions spectral analysis will show spectra with linewidths of about 100kHz.

This discussion shows that spectrum analysis with an S.A.W. analyser giving 50kHz resolution is well suited to this type of signal. When account is taken of the high radiation frequencies employed it is found that 50kHz corresponds to a line of sight velocity measurement to 0.25m/sec, i.e. 1MHz ≈ 5m/sec. Figure 12, shows a number of records obtained when different integration times are employed after spectral analysis. In most circumstances the spectrum obtained from a single batch of data (upper right) exhibits poor signal to noise. Integration over up to 600 batches of data show that in most circumstances the spectra remain complex but on some occassions linewidths in the region of 100kHz are obtained.

This system is now being developed to provide detailed information on windshear and aircraft wake vortices, both of which have been found to cause aircraft losses during landing.

6. CONCLUSIONS.

The design of a fast, highly linear spectrum analyser has shown that the S.A.W. technology can achieve high performance in application of the chirp transform algorithm. Application of the technique to pulse Doppler radar signal analysis and to laser anemometry show that the speed of response allows useful data to be extracted where alternative techniques would fail. Analysis of the unit characteristics shows that 40dB dynamic range and frequency determination to an accuracy of 0.01% can be achieved with total process time of 50μs.

7. ACKNOWLEDGEMENTS.

8. REFERENCES.

Alexander, S. J., Barbour, A., Ward, J., Alldritt, M., Jones, R., Vaughan, J. M., Watson, D., (1977), Proceedings of RADAR '77, IEE (London), to be published.

Cook, C. E. and Bernfield, M., (1967), Radar Signals, Academic Press, New York.

Hughes, A. J., Pike, E. R., (1973) Applied Optics, 12, 597.

Hughes, A. J., O'Shaughnessy, J., Pike, E. R., McPherson, A., Sparins, C., Clifton, R. H., (1972), Opto Electronics 4, 377.

Jacobs, W. W., Butler, M. B. N., Armstrong, G. A., (1976), Proceedings - First Military Electronics and Defense Expo, 153.

Roberts, R. B. G., Eames, R., McCaughan, D., Simon R., (1976), I.E.E.E. Trans SC-11, p. 100.

Sonnenschein, C. M., Horrigan, F. A., (1971), Applied Optics, 10, 1600.
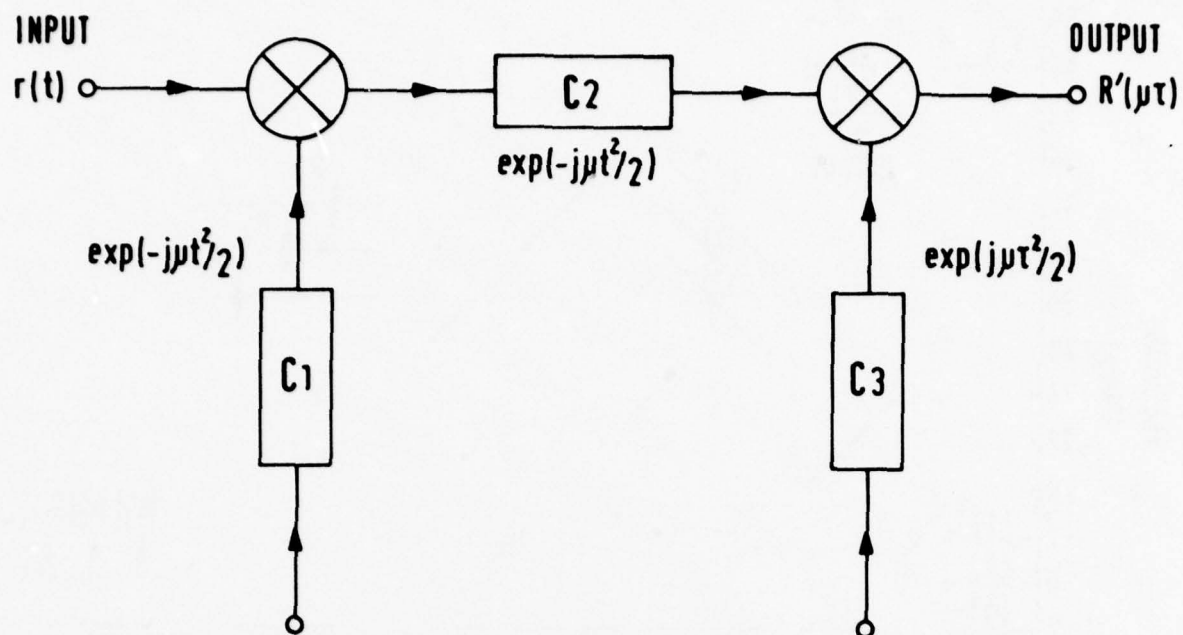
Figure 1.   Chirp - Z Transform implementation.
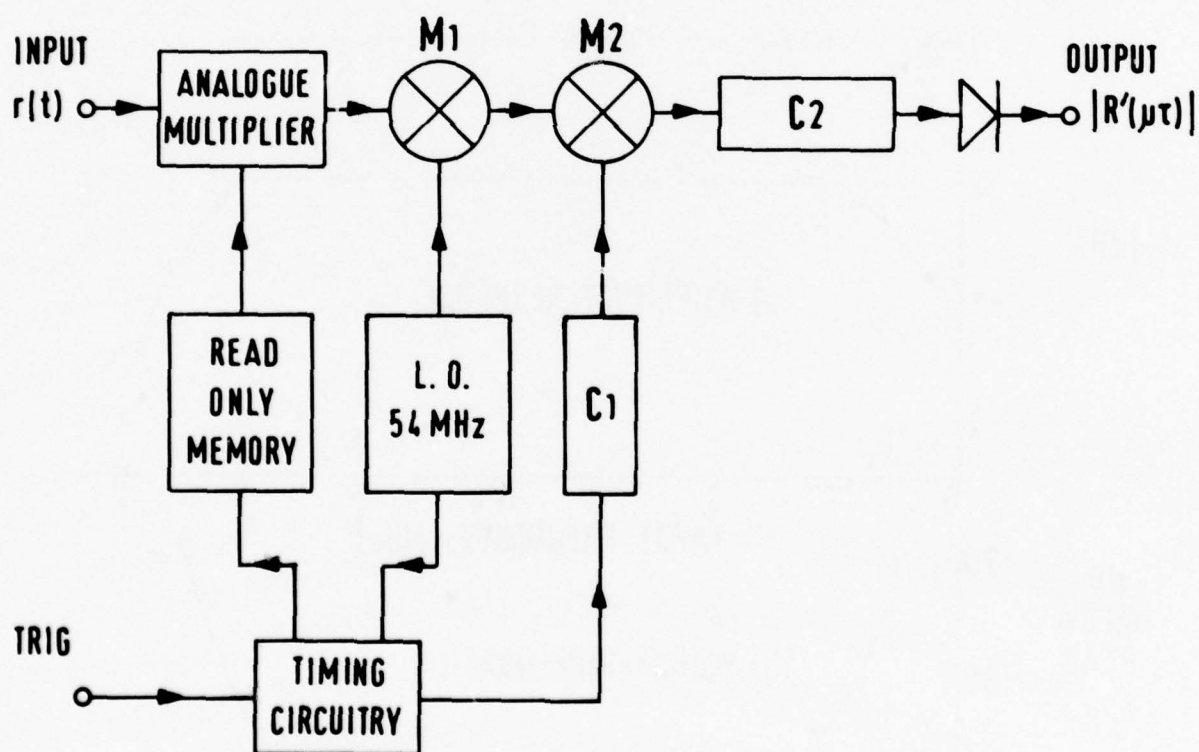
Figure 2.   Chirp Transform Spectrum Analyser.
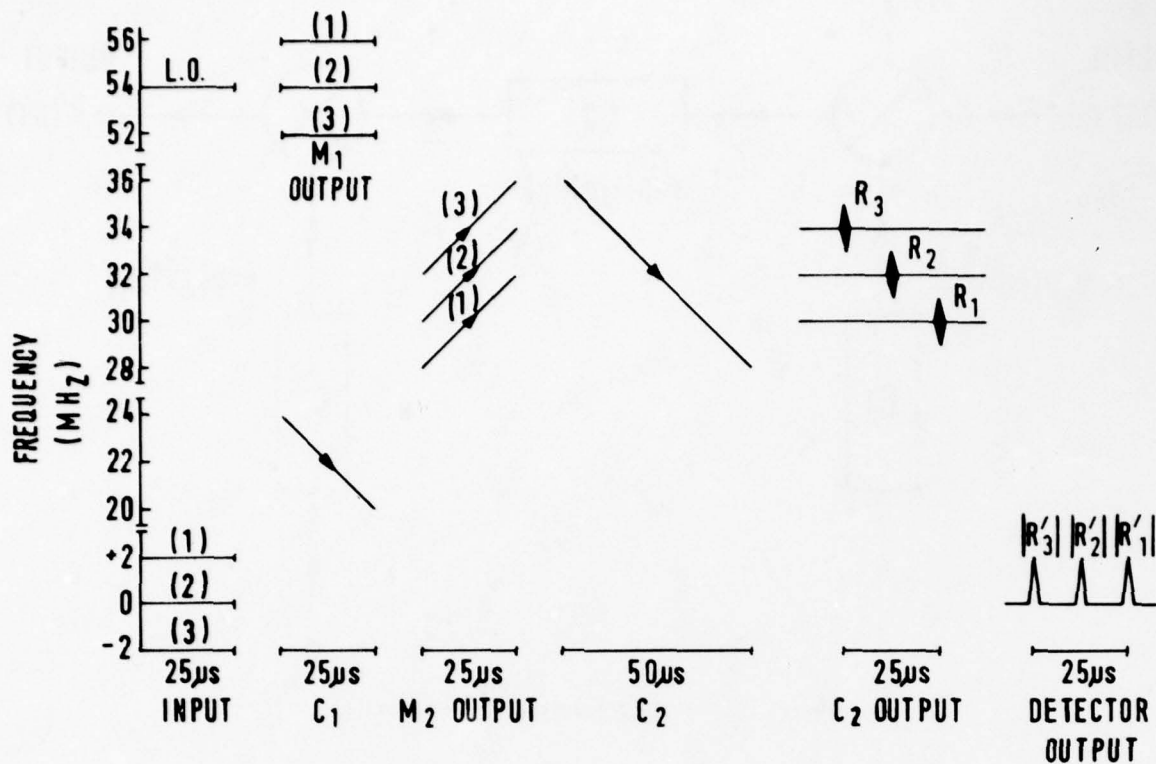
Figure 3.   Frequency - time diagram for the S.A.W. Spectrum Analyser.
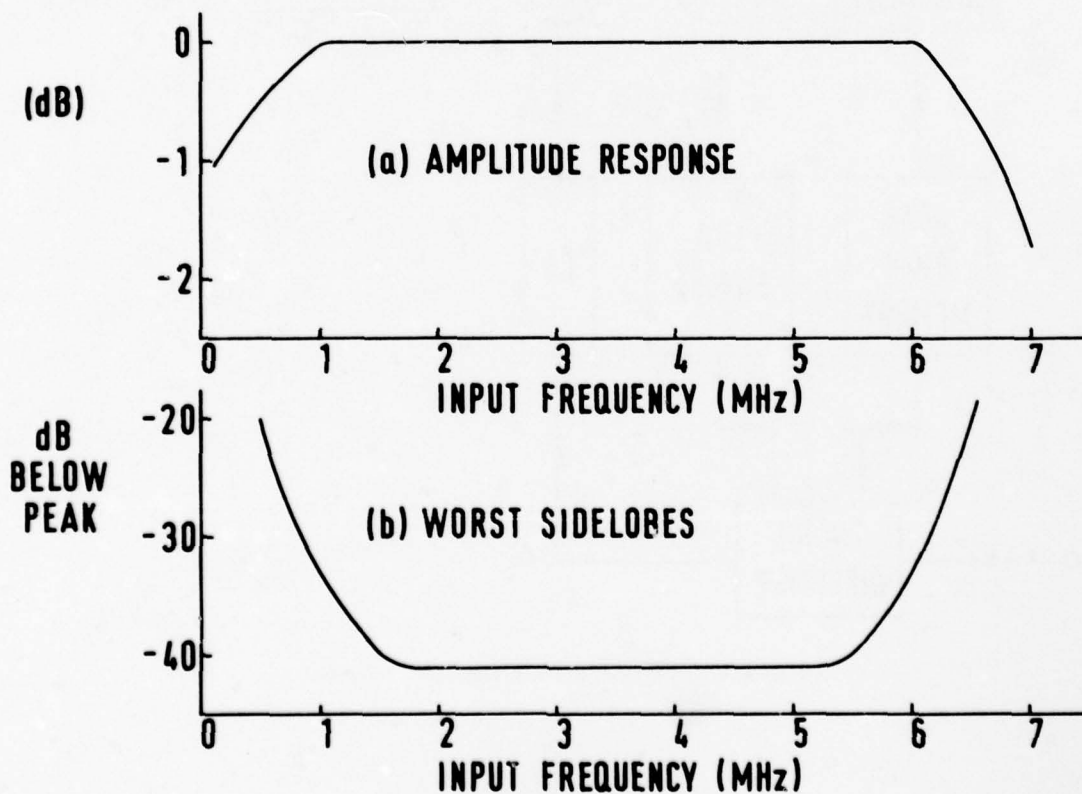
Figure 4.   Amplitude and sidelobe performance over an extended frequency range.
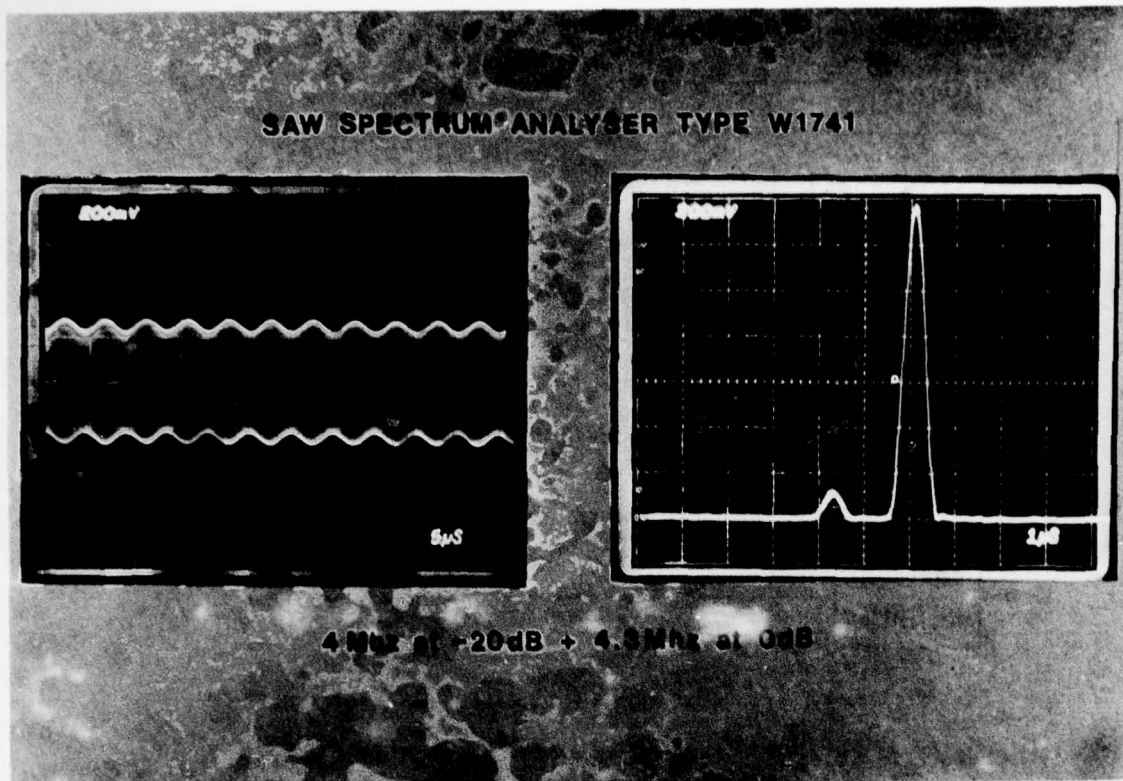
Figure 5.   Analyser output with 4.0MHz at -20dB and 4.3MHz at 0dB.
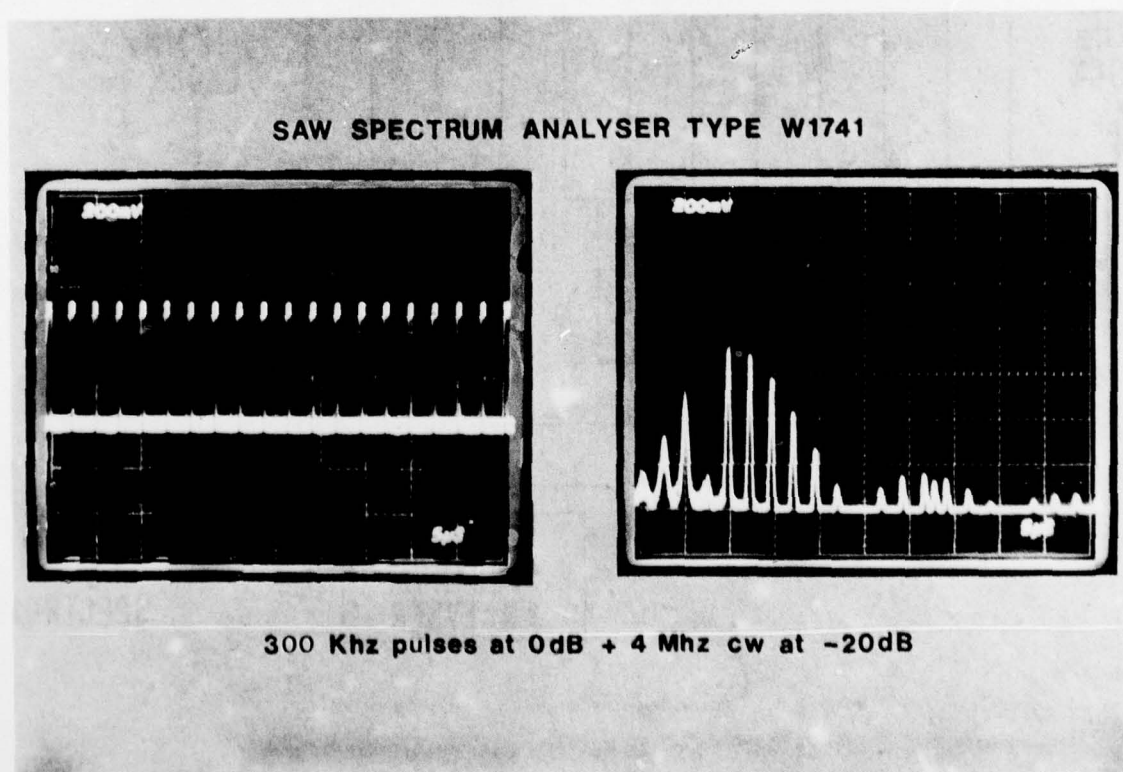
Figure 6.   Analyser output with 300kHz square wave at 0dB and 4.0MHz single tone at -20dB.
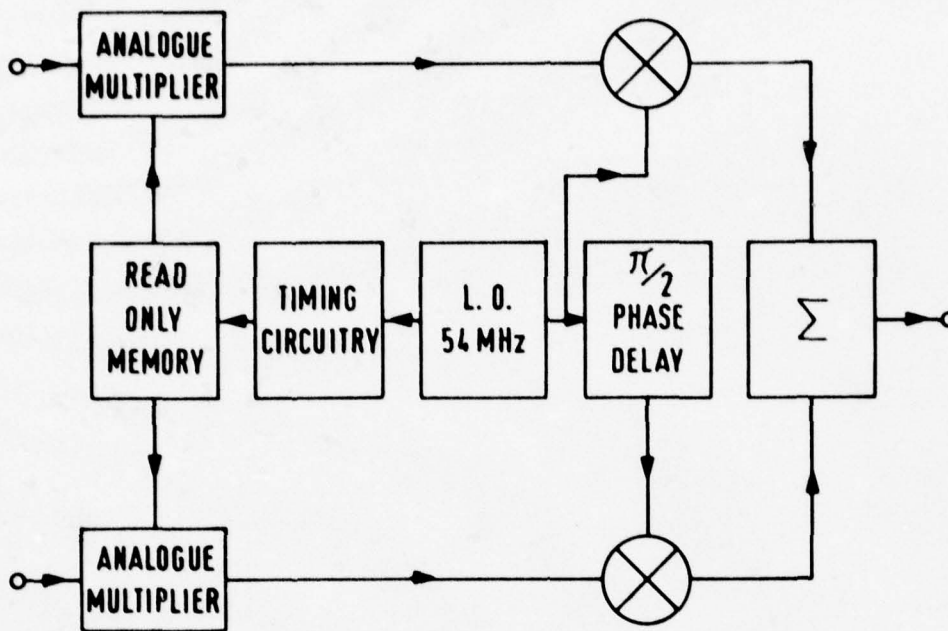
Figure 7. Input circuitry for double sideband operation.
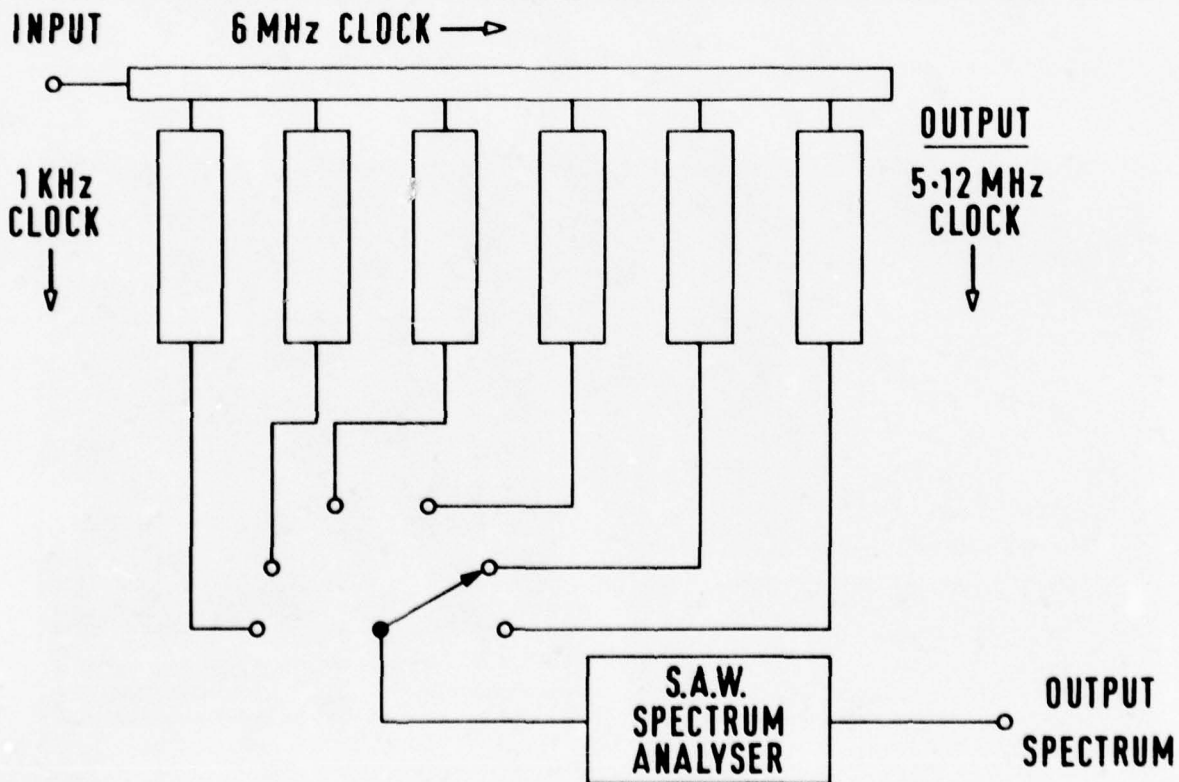
Figure 8. Pulse Doppler signal processing circuit.
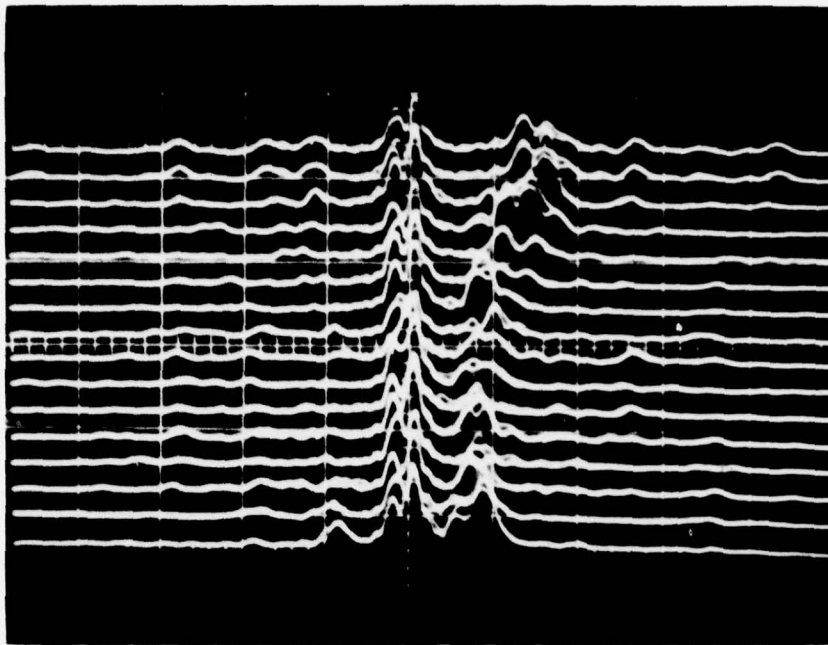
Figure 9. Spectra obtained in 16 range cells obtained near 6km. 6km range is shown in the lowest trace, cell separation 25m., display speed 2μs/div ≡ 3.44m/s per division.
© Copyright Controller HMBSO.
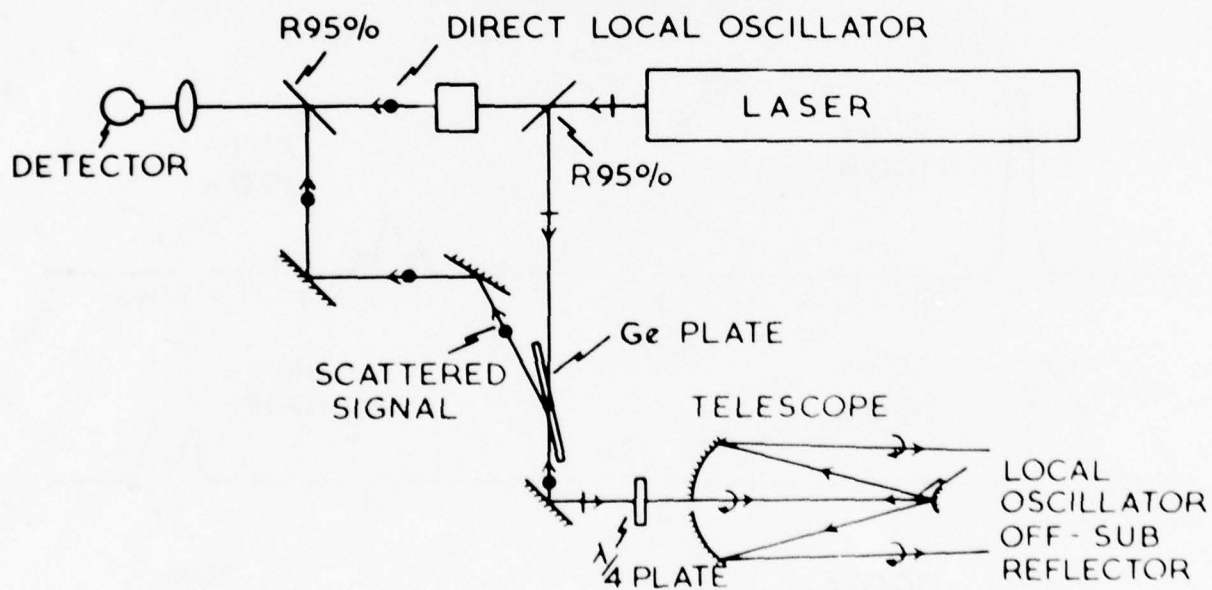
Figure 10. CO_2 laser anenometer optical system.
© Copyright Controller HMBSO.

$$r_o = \frac{2}{\pi} \lambda \frac{R}{D}$$

$$\Delta L = 2\frac{\pi r_o^2}{\lambda} = \frac{8}{\pi} \left(\frac{R}{D}\right)^2 \lambda$$

(a)

WEIGHTING
FUNCTION

(b)

↑50   ↑100 μs
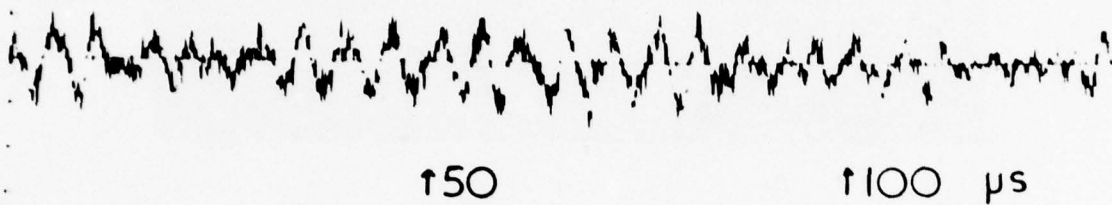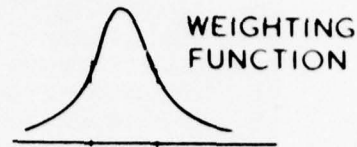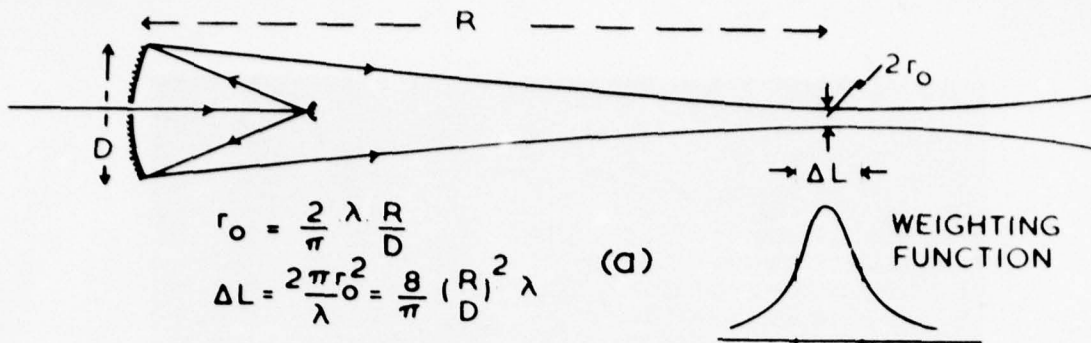
Figure 11.   Laser anemometer performance.
(a) Principal scattering region.
(b) Typical detector output.
© Copyright Controller HMBSO.
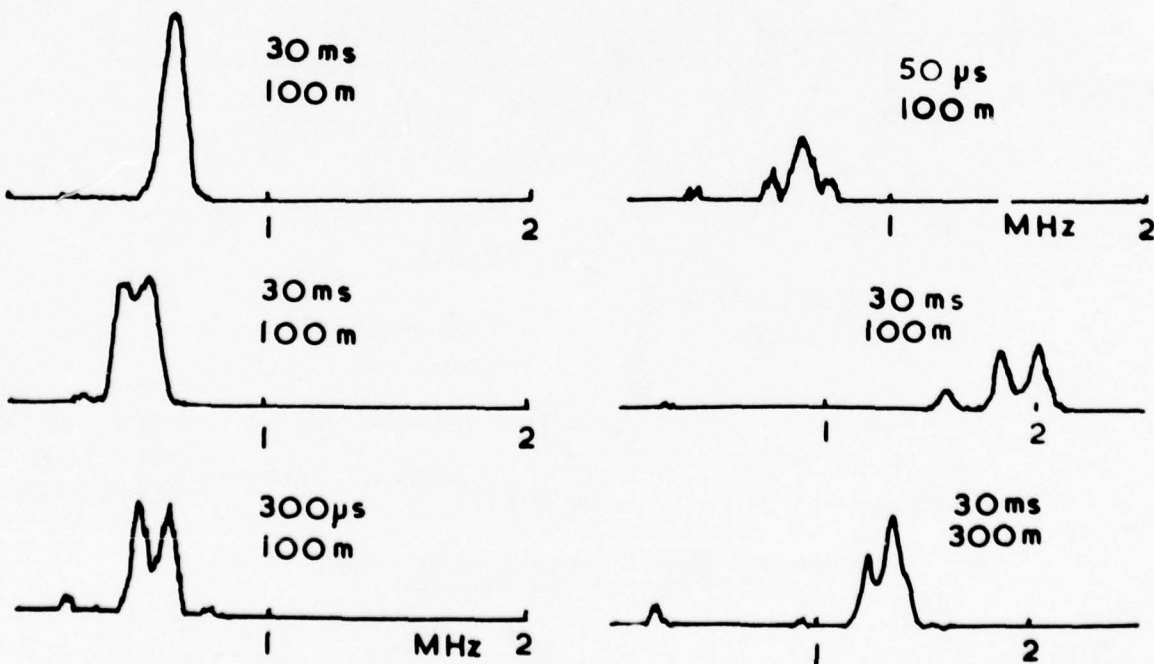
Figure 12.   Laser anemometer spectra obtained by integrating the output from
the Chirp Spectrum Analyser.
© Copyright Controller HMBSO.

THE CHIRP ≠ TRANSFORM WITH CCD AND SAW TECHNOLOGY

D. MacFall/J. Collins/W. Sciarretta/A. Cappon
Raytheon Company, Missile Systems Division
Bedford, MA

ABSTRACT

The combination of a Charge Coupled Device (CCD), as a sampling memory, and a Surface Acoustic Wave (SAW) dispersive delay line for pulse compression arithmetic, results in high throughput Fourier Transform signal processors that feature small size, low power, and low cost.

The Chirp ≠ Transform (C≠T) is presented as an IF process that makes use of time compression and chirp pulse compression. This allows sampled signals to be stored in CCD memory for milliseconds or more, and ≠ transformed by the chirp filter in microseconds. A C≠T spectrum analyzer therefore has extremely high throughput and the capability to process many channels with one processing element. Some applications require both magnitude and phase information. Configurations that achieve a linear complex output by removal of residual quadratic phase will be shown, resulting in side lobe levels that are independent of transform size.

A number of design areas have been addressed in order to achieve small, low cost, and reliable multichannel signal processing systems. These include the development of a 1K cell CCD Corner Turning Memory, a large time band-width product (1420) SAW chirp line on bismuth germanium oxide, and techniques for interfacing the analog C≠T processor with a digital computer.

1. INTRODUCTION

This paper will present a spectrum analysis technique with applications that is both faster and simpler than those currently employed in radar systems. This has been made possible by the development of dispersive Surface Acoustic Wave delay lines and Charge Coupled Devices. These technologies are used in a complementary manner to provide a signal processing capability which can dramatically reduce cost, size, and power consumption.

The CCD memory is used to sample and store the radar video output as analog charges. Once the required number of analog samples have been gathered, they are read out of memory in a time equal to the Expanded SAW pulsewidth. This analog memory is operated as a time compressing shift register to match radar dwell time to SAW chirp time. The time compression ratio scales long integration periods (milliseconds) to the SAW delay (microseconds). The chirp pulse expansion/compression system is used to perform the Fourier Transform in faster than real-time. This allows the chirp processing element to be time shared for the processing of many data channels. The time compressor memory is destructively read out to make room for new samples, since all processing is performed in one pass of the signal through the compression line.

The compression line is analogous to a comb-filter bank which is read out in natural ordered serial form. This is due to the fact that time delay in the compression SAW is proportional to frequency. Compressed pulses will appear at the output at a time proportional to the doppler frequency, with amplitude proportional to the input amplitude of each spectral component of the input time sequence of samples. The number of filters synthesized is proportional to the time bandwidth product of the compression system.

Frequency weighting of the expanded pulse, prior to multiplication by the time compressed output, is preferred to reduce time side lobes and thus improve the signal to clutter ratio for radar applications.

Modern system applications of the C≠T processor require that the output data be reported to a digital computer. A CCD memory can also be utilized to ease this interface. The analog memory can sample and store the filter amplitudes. Feedback is then utilized in the read cycle such that the only analog samples to be digitally converted are those that pass a given criteria, such as a threshold. This lowers the throughput such that an inexpensive slow A/D may be used.

2. THE CHIRP ≠ PROCESSOR

The Chirp ≠ Transform algorithm[1] can be implemented digitally or with analog transversal filters. The digital implementation is not as efficient as the FFT algorithm, however, the analog transversal filter approach (CCD or SAW) presents formidable competition. The hardware system described is implemented with SAW dispersive delay lines due to their available time-bandwidth products and inherent speed.

The basic Chirp ≠ Transform processor at IF is shown in Figure 1 using a weighted chirp premultiply waveform. This expansion chirp time is equal to the compressed dwell time and appears coincident for multiplication, as shown in Figure 2. The time compressed output is a "Box Car" waveform comprised of N samples of the radar receiver output S (t). The premultiply function is made up of a chirp (in this case a down chirp) and a weighting function W (t)

$$W (t) \, \text{Rect} \left(\frac{t}{KT} \right) \, \exp \left(j2\pi (f_o t - \frac{B}{T} \frac{t^2}{2}) \right) \tag{1}$$

B = Compression Chirp Bandwidth

T = Compression Chirp Time Duration

K = Gating Factor

f_o = IF Center Frequency of Compression SAW Delay Line

The result of this multiplication is:

$$S_i(t) \left[W(t) \ rect \left(\frac{t}{KT}\right) \ exp \left(j2\pi \left(f_o t - \frac{B}{T} \ \frac{t^2}{2}\right) \right) \right] \qquad (2)$$

The chirp pulse compressor is a SAW line with a chirp of the same slope $(-\frac{B}{T})$ but opposite sense to the premultiply function (here an up chirp).

$$rect \left(\frac{t}{T}\right) \ exp \left(j2\pi \left(f_o t + \frac{B}{T} \ \frac{t^2}{2}\right) \right) \qquad (3)$$

The pulse compression output can be determined analytically by a convolution of (2) and (3). This will yield a weighted compressed pulse at each of the frequencies contained in $S_i(t)$ at a proportional amplitude.

$$e_o(t) = exp \left(j2\pi \left(f_o t + \frac{B}{T} \ \frac{t^2}{2}\right) \right) \int_{-T/2}^{T/2} W(\lambda) \ S_i(\lambda) \ exp \left(-j2\pi \ \frac{B}{T} \ t\lambda\right) d\lambda \qquad (4)$$

The result in (4) is directly proportional to the weighted Fourier Transform of $S_i(t)$ in relative amplitude but with quadratic phase. When the power spectrum is desired, the compression output can be envelope detected to remove the IF carrier frequency with quadratic phase.

$$\sim \int W(\lambda) \ S_i(\lambda) \ exp \left(-j2\pi \ \frac{B}{T} \ t\lambda\right) d\lambda \qquad (5)$$

The expression in (5) is valid for large time-bandwidth products such that the residual quadratic phase is linear over the main lobe and near side lobes.[2] The expression above is essentially equal to the weighted Fourier Transform of $S_i(t)$, the receiver output.

In the frequency domain, each spectral component, f_d, of the radar output contained in $S_i(f)$ has translated the weighted expansion signal by that frequency as shown in Figure 3.

The transform pair that describe this operation is shown in Equations 6 and 7.

$$S_i(t) = rect \left(\frac{t}{KT}\right) cos \ 2\pi (f_o - f_d) \qquad (6)$$

$$S_i(f) = T \ sinc \ (\pi KTf) \ \delta \left(f_o - f_d \ \frac{T}{B}\right) \qquad (7)$$

The slope of the expansion chirp is equal but opposite in sense to the compression chirp to form a matched filter. The duration of the expansion chirp is smaller than the compression by a factor K, and therefore $0 < K < 1$.

A tradeoff is possible between the filter coverage $\left[B(1-k) \right]$ and frequency resolution $(\frac{1}{KT})$ by varying the gating parameter K (see Table 1). When $k = 0.5$ the maximum number of filters, $K(1-K)BT$, are synthesized. The expansion waveform can be generated by several methods. For minimum hardware, the value K should be determined and an expansion SAW line with dispersion KT should be designed with built-in weighting.

Figures 4a and 4b show the effect of frequency weighting. Here a gate and separate Hamming filter were used to demonstrate side lobe masking of a target 40 dB below a clutter signal for $K = 0.5$. Small targets can be detected much closer in doppler to larger ones with weighting, despite the main lobe broadening.

3.   THE COMPLEX SYSTEM

The system shown in Figure 1 accepts real samples and yields filter magnitudes. A complex system is required to determine the sense of vector rotation (approach versus recede doppler).

Figure 5 shows the full complex system with I and Q input and output as well as a quadratic phase cancellation postmultiply.

The I_o and Q_o outputs (Equations 8 and 9) retain the measure of the phase angle Θ of the input signal relative to the processing dwell.

$$I_o(t) = \frac{T}{2} \ cos \ \Theta_i \cdot \frac{sin \ \pi T (f_d + K\frac{B}{T})}{\pi T (f_d + K\frac{B}{T})} \qquad (8)$$

$$Q_o(t) = \frac{T}{2} \ sin \ \Theta_i \cdot \frac{sin \ \pi T (f_d + K\frac{B}{T})}{\pi T (f_d + K\frac{B}{T})} \qquad (9)$$

This system can be modified to do the following:

a. Complex in - Complex out (Figure 5)
b. Real in - Complex out
c. Complex in - Real out (Reference 5)
d. Real in - Real out (Figure 1)

The flexibility indicated here is attributed to the IF pulse compression arithmetic, since signal phase is preserved. This would not be the case if the process were at baseband, where a full complex process would require four compression lines (convolutions)[3] in order to retain phase. The system in Figure 5 requires one compression line since the I-Q precompression arithmetic is a single side band generator.

Many radar applications require only magnitude information from the spectrum analyzer for target detection purposes. However, the quadratic phase cancellation postcompressive multiply can prove beneficial particularly when small time bandwidth product pulse compression is used. The output is passed through a low pass filter to yield the Fourier Transform with side lobe performance approaching calculated levels.

The postmultiply provides a down conversion from IF to baseband and removes the effect of residual quadratic phase. Taylor and Hamming weighting use sets of paired echos that are added 180 degrees out of phase with the undesired side lobes for partial cancellation. With high time bandwidth products, the phase is essentially linear over the compressed pulse and the significant side lobes. The 180 degree phase cancellation is effective in this situation. However, for low time bandwidth products the quadratic phase is not linear over the significant side lobes. In this case there is a need to cancel the nonlinear phase to achieve the predicted side lobe suppression of the weighting function. With quadratic phase removed, weighted side lobe performance will be independent of the compression ratio as shown in Figure 6, for BT = 10.

4.   COMPARISON

At this point it is useful to compare the SAW CZT signal processor with other similar systems. D. D. Buss, et al,[3] made a comparison of the digital FFT and the all CCD CZT. These results show that a simple I^2L FFT produces filters at a 50 kHz rate, while the CCD CZT operates up to 5 MHz. The SAW CZT system shown here operates up to 20 MHz, since the CCD memory is a serial buffer rather than a complex split electrode arithmetic unit as in the all CCD approach.

For applications that require high speed processing, the digital FFT will require additional parallel arithmetic units. Thus, for high speed situations the CCD-SAW CZT will effect tremendous savings in hardware. Also, because of the memory configuration (Time Compression Corner Turner), and inherent SAW speed, one chirp Z processor can be time multiplexed among a number of input channels.

5.   CZT COMPUTER INTERFACE

New design techniques have been established and tested which reduce the complexity and cost of interfacing the CZT processor with a digital computer. In target acquisition and tracking radars, the information rate decreases as the data passes through the system. Data rates at the front end are high due to resolution and ambiguity requirements. The information rate is substantially reduced in passing through the signal processor. The information of interest usually pertains only to responses above threshold in the detection case and to the small amounts of data necessary to range, doppler, and angle track a target. Techniques have been developed to selectively pass on to the computer only those data that are meaningful. This has benefits in both reducing the A/D speed and the computer throughput.

The CZT to computer interface problem stems from the fact that the analog filter samples may have to be converted to digital form at high rates due to the throughput rate of the CZT, even though the average rate of desired information is low. This can occur when adjacent filters pass the threshold test and must be reported.

A solution to this problem is to incorporate an A/D converter with sufficient conversion speed to process the full output bandwidth. This, however, is expensive and would make the CZT approach less attractive when compared to the digital FFT.

CCD technology can be used to produce cost effective solutions for a problem. In this application the CCD functions fundamentally as a time expander. Two variations on this idea have been developed and tested, one for a detection channel in a ground based radar, and the other in the track channel of an airborne system.

The thresholded CCD buffered A/D was designed for use in a target detection channel as shown in Figure 7.

The CZT video output is sampled by a CCD at up to a 20 MHz write rate, while the read rate is variable depending on the threshold setting. As each analog filter sample is read from the serial CCD memory, its magnitude is compared to an externally supplied threshold voltage. The comparator output is time discriminated with the bit one A/D convert edge. At this instant of time, only valid data will permit a full 8-bit conversion. Then either the bit one convert edge or the end of conversion pulse is used to shift the CCD to the next sample and the next A/D conversion. The end of convert (EOC) pulse is used to interrupt the computer when a sample exceeds the threshold. The hybrid A/D controls the reading of CCD analog data samples as well as performing the conversion. A major reduction in A/D speed and computer throughput has been accomplished, since noise is discarded while only valid signals are fully digitized.

The second approach, for the track data, is shown in Figure 8. In this case, the amplitudes of a known group of filters is desired to provide inputs for the computer doppler and range tracking loops. The required filters are sampled and shifted into the CCD memory at a 5 MHz rate. The CCD clock is then lowered by more than an order of magnitude such that a 6 µs hybrid A/D converter can be used. It is necessary in this design to keep the dc offset introduced in the CCD time expander small such that tracking biases are minimized. This was accomplished by using two CCD registers in a full differential analog mode. Both CCD registers are in the same chirp and thus dark current induced offsets are well matched. An order of magnitude reduction in effective offset was obtained with the differential technique.

6.   C≯T TEST RESULTS

Two Chirp ≯ Transform signal processors have been constructed and tested for radar applications. One application, for missile borne equipment, stressed the importance of small size, low power, and low cost. The other application, in ground based radar, has as key features a large transform size and high throughput.

The system parameters of the missile processor are described in Table 2.

TABLE 2

Number of Channels	8
Number of Filter/Channel	128
Channel Process Time	50 μs
Filter Process Rate (throughput)	2. 5 MHz
Power Requirements	12 W
Size	2 modules (10 cm x 17 cm x 1. 25 cm)

A block diagram, Figure 10, shows the major components. This processor takes single input (real) data for each channel and delivers both real (magnitude) and complex outputs. A threshold is developed in the processor and the filter magnitudes are checked against it for target detections. When a target has been detected and verified, the complex outputs are used to provide the guidance computer with target amplitude and phase data.

The line addressable eight channel CCD memory samples and stores the receiver outputs for processing. When 455 samples of a given channel have been stored, the processing cycle for that channel begins. A voltage pulse, approximating an impulse, is impressed upon the expansion SAW delay line. This results in an expanded chirp signal with a duration of 25 μs. A 35 dB Taylor Weighting function was built in to the expansion SAW to reduce the frequency side lobes. Concurrently with the expanded chirp, the channel of the CCD memory is clocked out, as depicted in Figure 2. The two signals are multiplied, and the product enters the compression SAW delay line. The envelope of the compression SAW output is the Fourier Transform of the input sequence. This output is log magnitude detected to provide inputs for the target detection circuitry, (see Figure 9). A second expanded chirp is generated simultaneously with the compression line output. This second expanded chirp is used as a coherent LO for down converting the output to video while preserving phase information. This produces I and Q outputs for six midband filters. These twelve I and Q filter amplitudes are shifted at a 5 MHz rate into another CCD memory. They are subsequently read out at approximately 200 kHz for A/D conversion and computer interface. Each channel is processed in succession this way.

A photograph of the Chirp ≯ Transform modules is shown in Figure 11.

The parameters of the C≠T processor designed for the ground radar are described in Table 3.

TABLE 3

Number of Channels	32
Number of Filters/Channel	355
Channel Process Time	100 μs
Filter Process Rate (throughput)	3. 55 MHz
Coherent Integration Time	7. 6 ms
Filter Bandwidth	160 Hz
Power	20 W

The C≠T spectrum analyzer portion of this processor is similar in architecture to the unit previously described. The processor takes real inputs from the receiver and produces real outputs for the CCD buffered A/D converter.

The CCD memory consists of 32 line addressable 455 cell shift registers. They output time compressed data to the SAW processing element. The SAW delay line required to process this large transform has a time-bandwidth product of 1420 (described in the succeeding section).

Data are presented in Figures 12 and 13 to show the level of performance of the SAW processor. Figure 12 is the log detected output of the compression line. The input is a pure sinusoid at 2 MHz. The output shows the main compressed pulse with a second harmonic distortion term more than 40 dB down. Figure 13 shows the spectral analysis of a 700 kHz triangular wave in order to demonstrate multiple target capability. A triangular wave produces odd harmonics that reduce in voltage amplitude as the inverse square of the harmonic number. Table 4 lists the theoretical values.

The unit was designed to produce filters from zero frequency (on the left) to 7 MHz (tenth harmonic) which is the center of the scope display. Extra frequency coverage is provided due to the additional bandwidth of the compression SAW filter, which falls off slowly.

A complete 355 point transform is performed in only 100 μs, for a filter throughput rate of 3. 55 million filters per second. This is orders of magnitude faster than FFT processor.

The performance obtained to date with Chirp ≯ Transform processors indicate that the CCD time

TABLE 4
Test Triangle Wave Harmonics

Harmonic	Frequency (MHz)	Amplitude (-dB)
1	0. 7	0
3	2. 1	19
5	3. 5	28
7	4. 9	34
9	6. 3	38
11	7. 7	42
13	9. 1	45

compressor portion limits the overall dynamic range to 30 - 40 dB. This is due to the pattern noise experienced with memories such as the Fairchild CCD-321, Reticon SAM-64, and the Raytheon 1K buffer. The SAW pulse compression arithmetic section, however, has provided dynamic range in excess of 50 dB.

7.   DESIGN OF SAW DEVICES

The SAW delay lines utilized for the CZT processors are Reflective Array Compressors (RAC). This structure is illustrated in Figure 14. The grating structure consists of grooves which are ion beam etched into the substrate material. The relative spacing of the grooves controls the chirp rate and the groove depth variation along the substrate controls the amplitude function of the impulse response. A built-in weighting filter function can be obtained by appropriately controlling the groove depth profile.

A pair of SAW lines were developed for the 128 point processor on $LiNbO_3$ substrates. An up-chirp expansion line with weighting was packaged with a down-chirp flat-amplitude compression SAW. The use of the conjugate filter simplifies the design by eliminating the need for a spectrum inverter and external filter.

The design parameters for this pair of lines is displayed in Table 5.

TABLE 5

	Down-Chirp SAW	Up-Chirp SAW
Center Frequency	60 MHz	57. 5 MHz
Bandwidth	10 MHz	5 MHz
Chirp Rate	-0. 2 MHz/μs	+0. 2 MHz/μs
Insertion Loss	30 dB	30 dB
Weighting Function	Flat	35 dB Taylor
Dispersive Delay	50 μs	25 μs
Size (Pair)	12. 5 cm x 7. 5 cm x 1. 6 cm	

The SAW pair developed for the 355 point transform were constructed on substrates of bismuth germanium oxide (BGO). The design parameters are shown in Table 6. Because of the relatively long impulse response length (100 &s), BGO was chosen as the substrate material in view of its relatively slow SAW velocity. This permitted the substrate length to be only 4. 25 in., whereas equivalent $LiNbO_3$ devices would be nearly 9 in. long. Computed performance of the BGO SAW pair is shown in Figure 15.

TABLE 6

	Down-Chirp SAW	Up-Chirp SAW
Center Frequency	110 MHz	113. 55 MHz
Bandwidth (3 dB)	14. 2 MHz	7. 10 MHz
Chirp Rate	-0. 142 MHz/μs	+0. 142 MHz/μs
Insertion Loss	40 dB	35 dB
Weighting Function	Flat	35 dB Taylor
Impulse Response Length	100 μs	50 μs

8.   THE CCD MEMORY

Many benefits are derived by the incorporation of CCD technology into the SAW IF Chirp Z Transform Signal Processor. Both the SAW line and the CCD operate in the analog domain, and thus the use of high speed A/D converters can be avoided in the design of the spectrum analyzer. CCDs are especially advantageous when wide bandwidths are required because of their nonosecond sampling aperture. They also reduce memory size, since one analog storage location takes the place of approximately eight digital locations.

CCD memories are inherently low power devices because no static power is required in the storage element. The only power required is that to charge the clock capacitance, and to provide bias to the output

buffer circuit. The low power dissipation should also result in improved reliability as compared with digital T²L memories. In addition, low power density allows large amounts of storage on a single chip. This will further reduce the physical size of signal processors utilizing the technology.

8. 1   Device Considerations

Careful consideration must be given to the selection of a CCD memory for the time compressor memory. The parameters of the CCD will determine in large part the limits of application as well as the obtainable processor performance. The three governing CCD parameters are bandwidth, storage time, and dynamic range. The choice of a CCD involves determining whether surface or buried channel devices best match the processing requirements.

The bandwidth of a CCD is directly related to the efficiency of charge transfer from cell to cell. It is desirable to employ CCDs for the time compressor memory that have nE products less than 0. 1 at the read out frequency. The term nE is the product of the number of charge transfers from input to output (n), and the charge transfer inefficiency per transfer (E). The nE produce of 0. 1 translates to an attenuation of 1. 7 dB at the Nyquist frequency. [6] The smaller the nE product is for a given device, the more uniform the spectrum analyzer response will be over the frequency band.

C≠T processors require, for K = 0. 5 in Table 1, that the CCD bandwidth be equal to the SAW expansion line bandwidth. Current designs require bandwidths up to 5 MHz at clock rates as high as three times the bandwidth. The additional clock speed over the Nyquist rate may be required for oversampling (aliasing) considerations.

The buried channel CCD memories have a higher bandwidth and speed potential than does surface channel. The buried channel devices that have been used have 8 MHz of bandwidth at a 20 MHz clock rate. Available surface channel memories have been limited to less than 5 MHz bandwidth at a 10 MHz clock rate.

The storage time of a CCD is limited such that the peak to peak amplitude of the largest expected signal will not be clipped due to cell leakage current accumulation. Signal clipping will result in harmonic distortion terms that will severely limit the dynamic range of the processor. Also, because leakage current increases with increasing temperature, sufficient margin must be allowed for the accumulation of leakage current at the highest operating temperature. This margin has the effect of reducing the available dynamic range potential at lower operating temperatures.

The offset voltage generated in a time compressing memory has the form of a ramp function for uniform distribution of leakage current. This is due to the fact that each subsequent sample resides in the CCD memory a progressively smaller amount of time. The spectral power of this ramp appears in the lowest frequency filter along with IF multiplier feedthrough. This however, has not posed a limitation in system applications.

The surface channel CCD has superior leakage current characteristics and thus is useful in applications requiring long storage times. Data on existing Raytheon units show that 50 msec storage time is possible at 70°C. The commercial buried channel memories have been limited to less than 10 msec at 55°C.

The dynamic range of the CCD is the ratio of the largest signal that can be passed with acceptable distortion to the largest unwanted spectral component in the frequency band of interest. The largest signal level is set such that it does not clip at the highest operating temperature due to leakage currents. The unwanted components stem from CCD noise, CCD pattern signals, and pickup from colocated digital timing logic. The pattern signal is the set of output voltages that result when clocking the CCD with zero input signal. This variation can be caused by nonuniform cell leakage currents or by layout discontinuities such as in the serpentine structure. The voltage pulses that are produced are Fourier transformed by the SAW line pair. This results in a broad unwanted spectrum that masks otherwise detectable signals. This has been the major limitation in dynamic range with the commercial buried channel CCD memories. Experience has shown that 35 dB to 40 dB of dynamic range is achievable. This, however, is contrasted with the more than 50 dB of dynamic range available with the pulse compression arithmetic.

8. 2   Special Configurations

In order to further reduce size and complexity of the C≠T signal processor, more data must be stored in the CCD chip with an organization suited for pulse radar processing. This need has led to the concept of the 1K Corner Turning Memory. It is currently in development as a surface n channel device. A block diagram and photograph of the monolithic chip are shown in Figure 16.

Two blocks of 512 data cells are provided and operated in parallel from common clocks and address inputs. This forms two separate data channels, one for in-phase, and one for quadrature data samples. Each block of 512 cells is organized as eight 64 cell two-phase CCD registers with externally and independently addressable input and output select capability. The addressing is accomplished using two-phase dynamic digital shift registers integrated with the CCD.

The Corner Turning function is provided by multiplexing the eight CCD register inputs. This is accomplished by filling all input potential wells, and then spilling only the well of the CCD register that is addressed. Each register is addressed in succession. This is repeated 64 times to load the memory. The memory is read out by sequentially addressing each channel and providing 64 clock cycles.

The Corner Turner memory was designed to meet the performance goals as outlined in Table 7.

TABLE 7

Dynamic Range	40 dB
Insertion Loss	$\leq$ 0 dB
Input Sample Rate	15 MHz (Max)
Output Sample Rate	5 MHz (Max)
Power Dissipation	300 mW
Storage Time at 85°C (10 percent well filled)	3. 5 ms
Chip Size	120 x 150 mils

Fabrication is by a conventional poly/aluminum gate n channel CCD process using ion implantation to achieve threshold tailoring and two phase operation.

9.   SUMMARY

This paper has described an application where CCD and SAW technologies can be combined to achieve benefits in processor throughput, size, power, and cost. The CCD serves as both sampler and memory, and thus one monolithic network has replaced expensive high speed A/D converters and the associated digital memory. The passive SAW pulse compression Fourier Transform operates faster than conventional digital approaches and can be economically mass produced.

Experience has shown that improvements are necessary in CCD storage time and dynamic range to more fully realize the potential of the IF Chirp Z Transform processing technique.

REFERENCES

1. L. R. Rabiner, R. W. Shaffer, and C. M. Rader, June 1969, "The Chirp Z Transform Algorithm", IEEE Transactions on Audio and Electroacoustics, Vol. AV-17, No. 2, pp. 86-92.

2. J. R. Klauder, A. C. Price, S. Darlington, and W. J. Albershein, July 1960, "The Theory and Design of Chirp Radars", The Bell System Tech. Journal.

3. D. D. Buss, R. L. Veenkart, R. W. Brodersen, and C. R. Hewes, October 1975, "Comparison Between the CCD CzT and the Digital FFT", Proceedings 1975 International Conference on the Applications of Charge-Coupled Devices, San Diego, California.

4. J. D. Collins and W. A. Sciarretta, 25 March 1975, "Advanced Radar Signal Processing Techniques and Devices", The Reflector Boston Section of the IEEE, Vol. XXIV, No. 7, March 1, 1975, Technical Meeting of the Sonics and Ultrasonics, Information Theory, and Aerospace and Electronic Systems, Bedford, Massachusetts.

5. J. D. Collins, W. A. Sciarretta, D. MacFall, and M. D. Schulz, September 29, 1976, "Signal Processing with CCD and SAW Technologies", Proceedings 1976 Ultrasonics Symposium IEEE 76CH1120-SSU.

6. Sequin and Tompsett, Charge Transfer Devices, 1975, Academic Press.
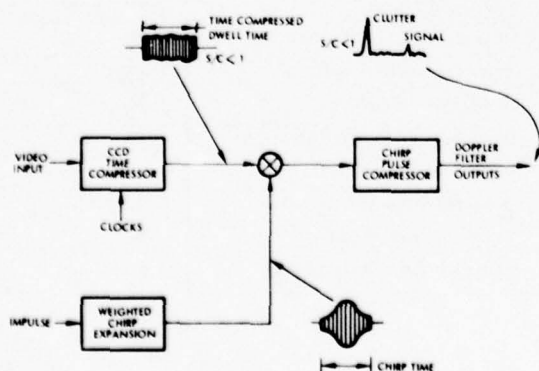
Figure 1 - The Basic Implementation of the SAW Chirp Z-Transform
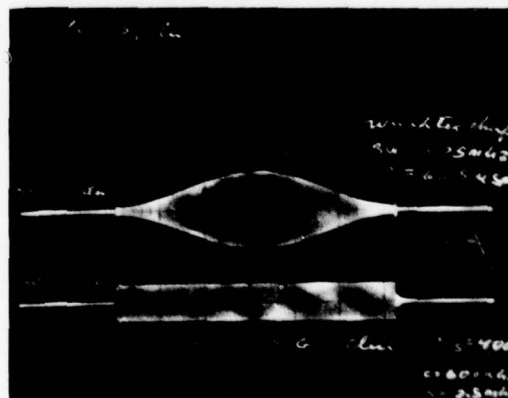
Figure 2 - Multiplication Factors: Weighted Chirp and Gated Video
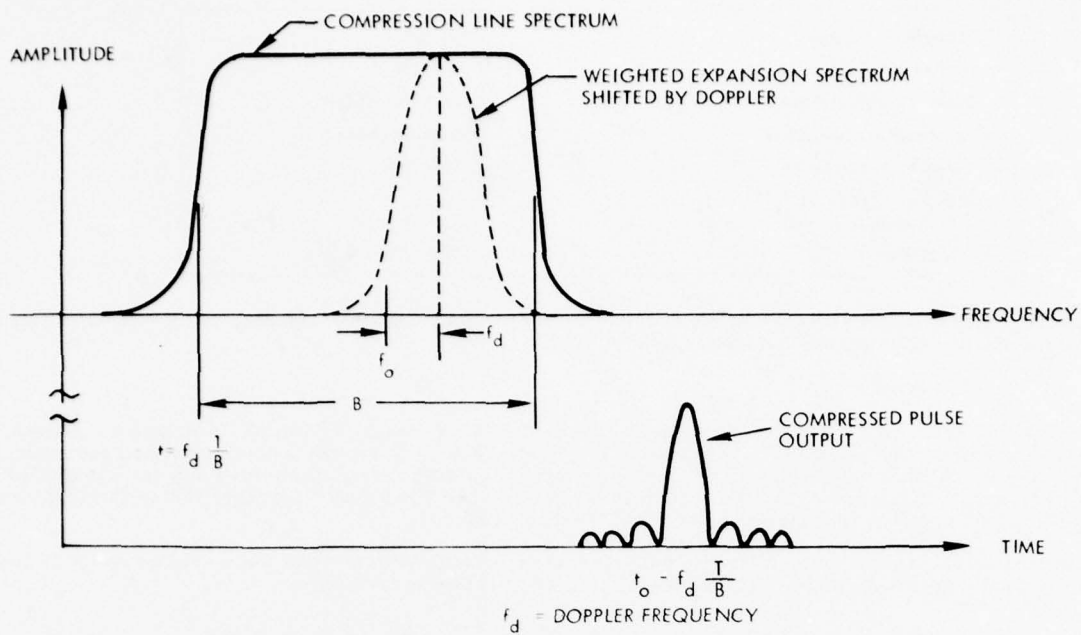
Figure 3 - Time-Frequency Relationship of Processor Output

a) Weighted Output (k = 1/2)
with T - 62.5 μsec

b) Unweighted Output (k - 1/2)
with T = 62.5 μsec

Figures 4a, b, - Clutter Signal at 600 kHz
and Target at 2.5 MHz; 40 dB Down
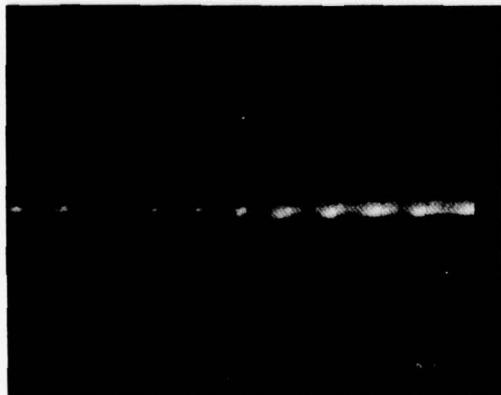
TABLE 1
Chirp FT Parameters
(Unweighted)

Parameter	Chirp FT
Sample Time	$\dfrac{kT}{N}$
Sample Rate	$\dfrac{N}{kT}$
Dwell Time	kT
Filter Bandwidth (Unweighted)	$\dfrac{1}{kT}$
Filter Coverage	$\pm \dfrac{B}{2}\ (1-k)$
Number of Filters	$k(1-k)\,BT$
Time Compression Ratio	$\dfrac{Nt}{kT}$
Pulse Compression Ratio	$k^2 BT$
Compressed Pulsewidth	$\dfrac{1}{kB}$
Dispersive Bandwidth (Gated)	kB
Dispersive Time (Gated)	kT
Pulse Expansion Gating Factor	$0 > k > 1$
Number of Input Samples	N
Processor Input Sample Rate	$\dfrac{1}{t}$

Figure 5 - Complex SAW Chirp Z - System
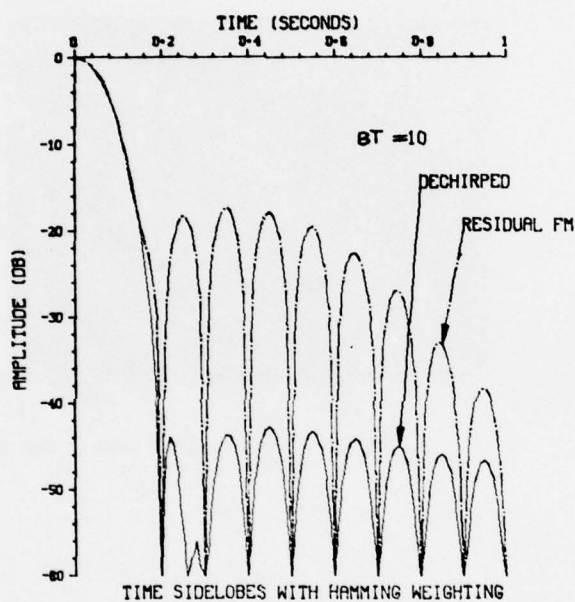
Figure 6 - Computer Simulation Shows Effect of Residual Quadratic Phase on BT = 10
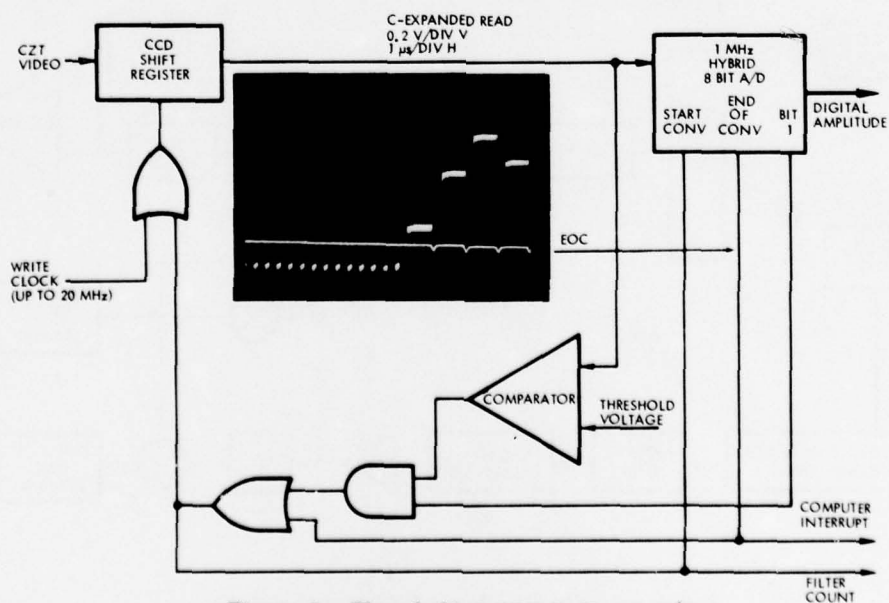
Figure 7 - Thresholded CCD Buffered A/D
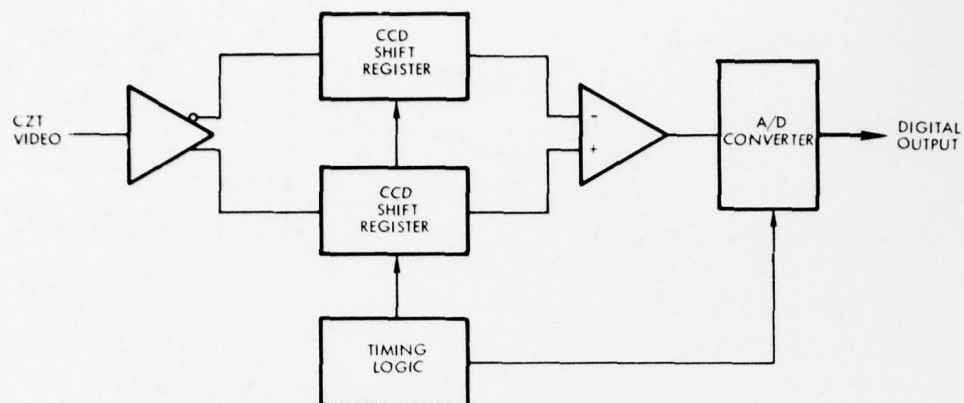
Figure 8 - Differential CCD Time Expander
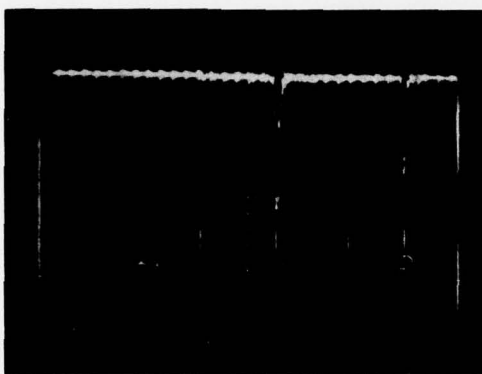
Figure 9 - Spectrum Analyzer Magnitude Output
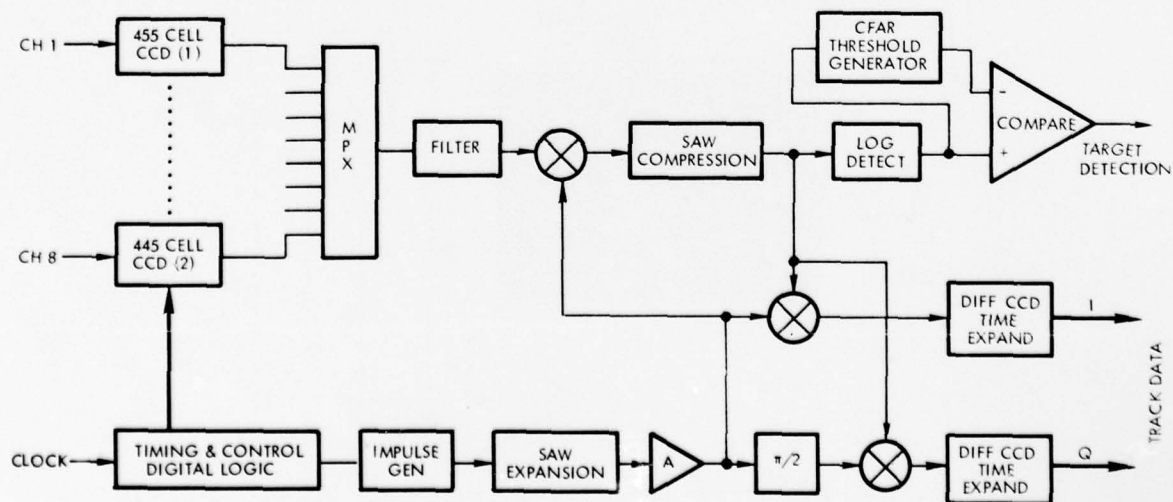
Figure 10 - Missile Borne DZT Signal Processor
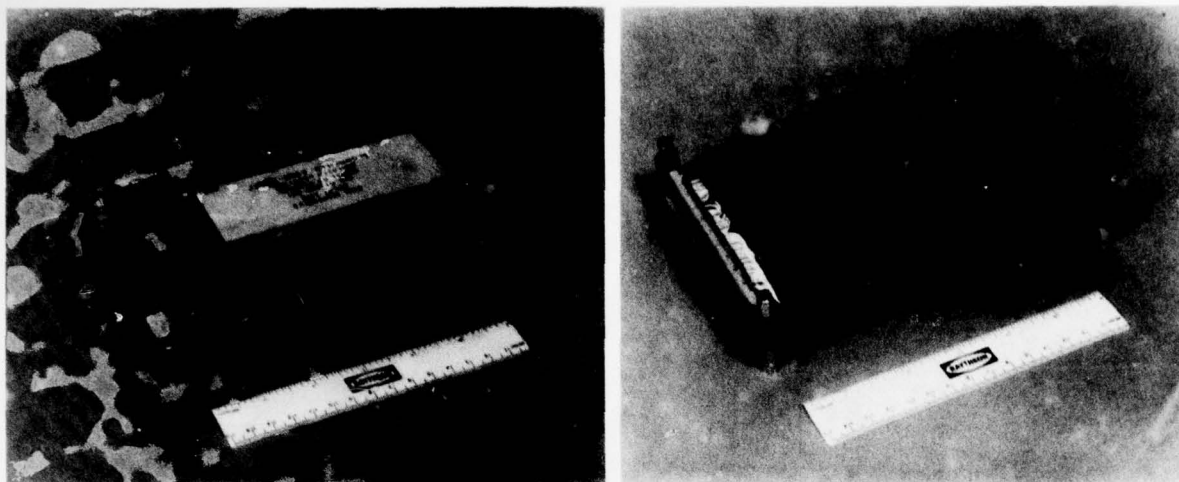
Figure 11 - CZT Missile Processor
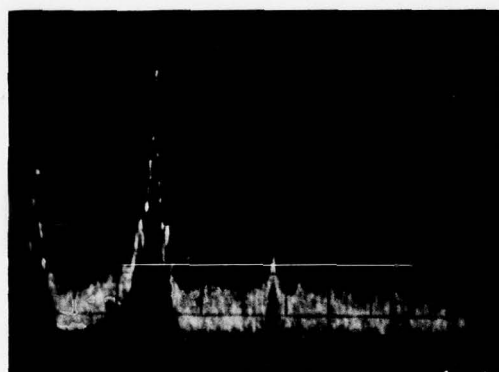
VERTICAL 200 mv/div ≈ 10 dB/div
HORIZONTAL 5 µs/div ≈ 700 kHz/div
          -38 dBm input signal

VERTICAL 200 mv/div ≈ 10 dB/div
HORIZONTAL 10 µs/div ≈ 1.4 MHz/div

Figure 12 - 2 MHz Sine Wave Video

Figure 13 - 700 kHz Triangular Wave Video
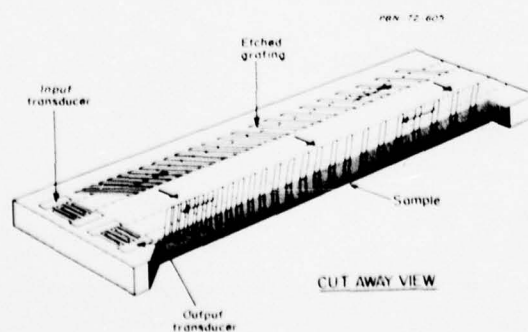
Figure 14 - Reflective Array Compressor
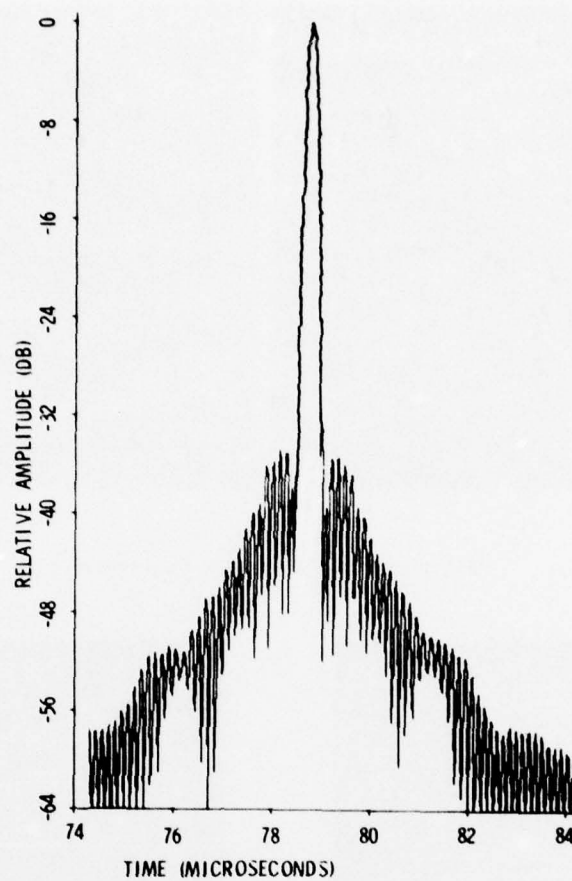
Figure 15 - Expected Time Domain Response of
BGO Device for 355 Point Transform System
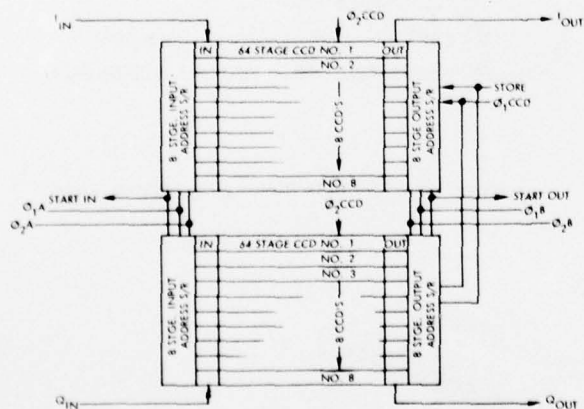
Figure 16 - 1K Cell Corner Turner Block Diagram and Chip Photo

DISCUSSION

Roberts

1 — What level of spectral sidelobes were actually achieved?

2 — When using CCD time compression stores for I and Q data, do you find it a problem to match the two channels well enough to allow adequate image rejection?

Author's Reply

1 — The spectral sidelobes achieved in the missile processor were 32 db below the peak of the compressed pulse.

2 — We have experienced a problem in matching the responses across the range channel of the processor. It is necessary at this time to select CCD devices to achieve matching to within one db. I am sure device selection would have to be done to achieve a sideband suppression of more than 35 db.

SPECTRAL ANALYSIS USING THE CCD CHIRP Z-TRANSFORM

W. L. Eversole, D. J. Mayer, P. W. Bosshart
M. de Wit, C. R. Hewes, and D. D. Buss
Texas Instruments Incorporated
Dallas, Texas

SUMMARY

The CCD analog transversal filter is a tremendously cost-effective component in terms of its simplicity compared to equivalent digital hardware. In view of this, the chirp z-transform (CZT) algorithm for performing spectral analysis is ideally suited to CCD implementation because, in this algorithm, the bulk of the computation is performed in a transversal filter. The CCD CZT has some performance limitations relative to the digital fast Fourier transform (FFT), and for this reason, it is not applicable to all signal processing systems. However, for those applications which fall within the CCD performance capabilities, the CZT offers significant potential cost savings over the digital FFT. This paper begins with a review of CCD CZT spectral analysis techniques. Results are then presented on

- A 500-point CCD CZT implemented with CCD filters and discrete components
- A 32-point completely integrated CCD CZT IC.

The paper concludes with a discussion of CCD CZT applications.

1. INTRODUCTION

Charge-coupled device (CCD) analog transversal filters can be used to perform the discrete Fourier transform (DFT) on electrical signals using an algorithm called the chirp z-transform (CZT). (Whitehouse, Speiser, Means, 1973)(Rabiner, Schafer, Rader, 1969) The CCD CZT has been demonstrated, (Brodersen, Hewes, Buss, 1976) and it has been proposed as a replacement for the digital fast Fourier transform (FFT), in a number of signal processing applications. (Buss, Veenkant, Brodersen, Hewes, 1975) It is the purpose of this paper to review CCD CZT spectral analysis techniques and to present results on state-of-the-art CCD CZT technology.

Section 2 reviews the CZT algorithm, and discusses the advantages of CCD implementation. The sliding CZT which is useful in many spectral analysis applications is discussed, and the performance limitations of the CZT are reviewed.

Section 3 presents the design and evaluation results on two CCD CZT developmental systems

- A 500-point sliding CZT which is implemented with two CCD filter ICs together with discrete circuitry.
- A completely integrated 32-point CCD CZT IC.

At the time of the writing of this manuscript, evaluation of the 32-point CZT unit is not complete and is omitted from the text. Evaluation will be completed by October 1977, however, and detailed performance results of the 32-point CZT IC will be presented orally at the conference.

Section 4 concludes with a discussion of the application of the CCD CZT to radar signal processing, video bandwidth reduction, speech processing and other systems requiring spectral analysis.

2. CCD CZT SPECTRAL ANALYSIS

The Fast Fourier Transform (FFT) algorithm is currently the most widely used technique for implementing spectral analysis. It is important because it requires only $N/2 \log_2 N$ complex multiply operations to perform an N-point Discrete Fourier Transform (DFT) as compared with N^2 complex multiply operations required to directly implement the DFT formula

$$F_k = \sum_{n=0}^{N-1} f_n e^{-i2\pi nk/N} \tag{1}$$

From the standpoint of minimizing the number of digital operations required to perform the DFT, the FFT algorithm is optimum. However, in determining the optimum algorithm for implementation with analog CCD's, a whole new set of ground rules exists. It is no longer important to minimize multiplications, because CCD transversal filters can be built which perform a large number of multiplications simultaneously in real time. Consequently, for CCD implementation, algorithms should be selected in which the bulk of the computation is performed by a transversal filter. Two such algorithms have been identified for cost-effective CCD implementation; the CZT and the prime transform. (Radar, 1968) Only the CZT is discussed in this paper.

2.1 CZT Algorithm

The CZT gets its name from the fact that it can be implemented by (1) premultiplying the time signal with a chirp (linear FM) waveform, (2) filtering in a chirp convolution filter, and (3) postmultiplying with a chirp waveform. When implemented digitally, the CZT has no clear cut advantages over the conventional FFT algorithm. However, the CZT lends itself naturally to implementation with CCD transversal filters.

Starting with the definition of the DFT given in equation (1), and using the substitution

$$2nk = n^2 + k^2 - (n - k)^2 \tag{2}$$

the following equation results:

$$F_k = e^{-i\pi k^2/N} \cdot \sum_{n=0}^{N-1} \left(f_n e^{-i\pi n^2/N} \right) e^{i\pi(k-n)^2/N} \tag{3}$$

This equation has been factored to emphasize the three operations which make up the CZT algorithm. It is illustrated in Figure 1.

To implement the conventional N-point CZT, the CCD filters are chirp filters of length 2N-1 which chirp from $-f_c$ to $+f_c$, and the premultiply waveform has a time duration N/f_c and chirps from zero to $-f_c$. A physical interpretation in terms of correlation of the input chirp with the filter is given in Figure 2. When the input signal has zero frequency, the product with the premultiply chirp results in an input waveform to the filter which chirps from 0 to $-f_c$. The samples corresponding to frequencies near $f = 0$ are clocked into the filter first, and those near $f = -f_c$ are clocked in last. This sequence of samples results in a correlation peak at $t = t_o$, when the product waveform has been clocked into the first half of the filter. When the input frequency is $f_1 \neq 0$, the product with the premultiply chirp results in an input to the filter which chirps from f_1 to $-f_c + f_1$. The input waveform (V_{in} x chirp) in Figure 2 corresponds to an input signal at a frequency f_1 at time $t = t_o$. This waveform is shifted to the right as t increases resulting in a correlation peak at t_1. The shift in time relative to the dc correlation peak is

$$t_1 - t_o = \frac{N}{f_c^2} f_1 \tag{4}$$

In this way, the time axis of the output is calibrated in frequency. The postmultiply is needed to obtain the proper phase of the DFT coefficients and can be omitted when phase is not required.

Several undesirable features of this implementation become apparent from the above description. The output must be blanked during the loading of the chirp into the filter, and the input must be set to zero during the calculation of the coefficients. Also undesirable is the inefficient use of the CCDs since only half of the CCD filter has useful information at any point in time.

2.2 Sliding CZT

For DFT applications, such as video bandwidth reduction, (Means, Whitehouse, Speiser, 1974) in which the transform is to be inverted to regain the original signals, the CZT is performed in the way described above. However, in many applications, it is desired to obtain a measure of the spectral energy density of a quasi-periodic waveform or a quasi-stationary random process. For these applications, the phase of the transform is not important. In addition, the true DFT is not performed, but the input data are apodized or windowed by an appropriate windowing function w_n to suppress frequency sidelobes in the spectrum. (Rabiner, Gold, 1975) The desired result is

$$|F_k|^2 = \left| \sum_{n=0}^{N-1} f_n w_n e^{-i2\pi nk/N} \right|^2 . \tag{5}$$

Using the substitution of equation (2) in equation (5) gives

$$|F_k|^2 = \left| \sum_{n=0}^{N-1} \left(f_n w_n e^{-i\pi n^2/N} \right) \cdot e^{i\pi(k-n)^2/N} \right|^2 . \tag{6}$$

In this case, simplification of the CZT algorithm results from two observations: (1) The postmultiply operation can be eliminated and (2) The sliding CZT can be used.

The sliding DFT is defined in this paper to be

$$F_k^s = \sum_{n=k}^{k+N-1} f_n e^{-i2\pi nk/N} \tag{7}$$

and it gives a windowed power density spectrum

$$|F_k^s|^2 = \left| \sum_{n=0}^{N-1} f_{n+k} w_n e^{-i2\pi nk/N} \right|^2 \tag{8}$$

$$= \left| \sum_{n=0}^{N-1} \left(f_{n+k} w_n e^{-i\pi n^2/N} \right) \cdot e^{i\pi(k-n)^2/N} \right|^2 \tag{9}$$

Comparison of equations (8) and (9) with equations (5) and (6) indicates that the sliding CZT differs from the conventional CZT in that the sliding CZT indexes the data each time a spectral component is calculated. For a periodic waveform, indexing results in a phase factor which does not affect the result, and for a stationary random signal, the time record is different for each spectral component but stationarity insures that the result is unaffected. For these two classes of signal the sliding CZT gives the same result as the conventional CZT.

Figure 3 gives a pictorial comparison between the conventional CZT and the sliding CZT for the simple case of a 3-point transform. With the conventional CZT, all three Fourier coefficients F_0, F_1, F_2 are calculated using the first three time samples f_0, f_1, f_2. These coefficients are being calculated by the filter during the next three clock periods, so that time samples $f_3 - f_5$ must be blanked. Then the cycle repeats as shown in Figure 3a. Using the sliding CZT, F_0^s is calculated on the sample record f_0, f_1, f_2 as before, but F_1^s is calculated on the sample record f_1, f_2, f_3, F_2^s on the record f_2, f_3, f_4, and the next F_0^s computation is made of the sample record f_3, f_4, f_5. The sample record is continually updated by replacing the oldest sample with a new one. The above description shows that N Fourier coefficients are obtained for N time samples (100% duty cycle).

The advantages of the sliding CZT are (1) For an N-point transform, N-stage filters are required which chirp through a bandwidth f_c ($-f_c/2$ to $+f_c/2$ for example). (2) No blanking is required. The filters operate with 100% duty cycle; i.e., one spectral component out for each time sample in. (3) Windowing can be achieved by weighting the chirp impulse response of the filter with the desired window function. (4) The degradation due to imperfect charge transfer efficiency is less for the sliding CZT than for the conventional CZT.

2.3   Accuracy of CCD CZT

The sources of error in a CCD CZT are (1) thermal noise, (2) quantization of the pre- and postmultiply chirp waveforms, (3) weighting coefficient error in the CCD transversal filters, and (4) CTE. When the criterion of rms error to rms signal is applied, imperfect CTE generates large errors, because the errors add coherently. Because of this fact, however, CTE effects can be treated as a resolution degradation and not as "random" error.

Thermal noise is analogous to input quantization in a digital FFT because it generates an error which is independent of signal size. Assuming the rms noise referred to the input is 60 dB below the maximum peak signal, the equivalent quantization accuracy is 8 bits plus sign. At higher signal levels thermal noise, like input quantization noise in a digital FFT, is dominated by signal dependent errors.

Errors which result from the 8-bit quantization of the chirp signals dominate the rms error of the 500-point CZT and given rise to an rms error to rms signal of about .003. (Brodersen, Hewes, Buss, 1976)

Weighting coefficient error arises from a number of sources, but let us assume as a model, that the placement of the gap in the split electrodes is quantized in steps of δ during photomask fabrication. δ can be made as small as 10 μin, and the channel width W is typically 5 mil giving $\delta/W = .002$. This is equivalent to quantizing the weighting coefficients to 8 bits plus sign and is analogous to twiddle factor quantization in the FFT. The rms error to rms signal which would result for a 500-point CCD CZT is on the order of .0008. (Brodersen, Hewes, Buss, 1976)

The two error sources discussed above are independent of the number of points N in the transform and are indicated by the dashed line in Figure 4. Also shown in Figure 4 are the results of a computer simulation of error using randomly generated errors.

For comparison with the digital FFT, a block floating point truncation algorithm is assumed.

The most important source of error in this type of FFT is usually overflow and round-off of data words during butterfly computation. If the data words are carried with "b" bits plus sign, an upper bound on error in a block floating point machine can be determined assuming overflow occurs at every stage. The result is (Welch, 1969)

$$\Delta_B = .3 \sqrt{8} \ N^{\frac{1}{2}} \ 2^{-b} \tag{10}$$

If the twiddle factors are quantized to the same accuracy as the data words, Δ_B dominates FFT error. Although equation (10) does not contain the input signal size explicitly. Δ_B does scale in a general way with signal because for smaller signals, overflow does not occur at every stage. The dependence of Δ_B on the length of the transform indicates that higher accuracy (larger b) is required for longer transforms.

For fixed b, Δ_B increases like $\sqrt{N}$, and it is plotted in Figure 4 for the case b = 13. Figure 4 shows that a 512-point, block floating point FFT with 14-bit internal arithmetic is comparable in terms of rms error to the CCD CZT.

The performance limitations of the CCD CZT and the digital FFT are summarized in Table 1.

3.   CCD CZT DESIGN AND PERFORMANCE

In this section, some previously published results on a 500-point sliding CZT system are reviewed, (Brodersen, Hewes, Buss, 1976) and the design of a completely integrated 32-point CZT is presented. At the time of this writing, evaluation of the 32-point CZT IC is not complete. Evaluation will be completed by October 1977, however, and detailed performance results will be persented orally at the conference.

3.1   Performance of the 500-Point CCD CZT

The block diagram for obtaining the spectral density using the sliding CZT is shown in Figure 5. The rectangles represent CCD filters having impulse responses $w_n \cdot \cos \pi n^2/N$ and $w_n \cdot \sin \pi n^2/N$, $-N/2 < n < N/2 - 1$. This system has been implemented using 500-stage CCD filters. If a windowed transform is required, the desired window function w_n is coded into the photomask. Systems have been demonstrated without windowing and with Hamming windowing. (Brodersen, Hewes, Buss, 1976)

The experimental implementation of Figure 5 utilized two CCD ICs each containing two 500-stage filters. The premultiply chirps were stored with 8-bit precision in ROMs. Multiplication was performed in discrete multiplying digital to analog converters (MDACs), and the squaring operation was performed in bipolar analog multipliers. This experimental demonstration required 33 IC packages in addition to the two CCD packages and 16 discrete MOSFETs.

The operation of this system is illustrated in Figures 6 and 7. Figure 6 shows the response of the system clocked at 20 kHz to sinusoidal input signals. Figure 7 shows the response of the system clocked at 25 kHz to a 200 Hz square wave. The output spectrum shows odd harmonics which decrease in amplitude approximately as $1/n^2$. The effect of imperfect CTE is visible in Figure 7c as indicated by the arrow.

The system of Figure 5 is limited in dynamic range by the squaring amplifiers, which are implemented with analog multipliers. The multipliers have output dynamic range of approximately 80 dB, thereby limiting the input dynamic range to only 40 dB. The overall system up to the squaring multipliers has approximately 70 dB dynamic range, so the squaring multipliers are the weak link that limits dynamic range. To circumvent this problem, an improved magnitude circuit was implemented using bipolar amplifiers, and the output was applied to a log amplifier. The dynamic range improvement is illustrated in Figure 8. In the top photograph the maximum signal is applied to a CZT unit with Hamming windowing. With the peak at 0 dB the sidelobes are at -40 dB. In the middle photograph the input sinusoid has been attenuated by 20 dB, and in the bottom photograph it has been attenuated by 40 dB. The noise level in the bottom photograph is around -60 dB. This illustrates clearly that the CCD CZT is capable of 60 dB dynamic range.

3.2   CCD CZT IC

As seen in equation 3 and Figure 1, the CZT algorithm is made up of three operations: premultiply, convolution and postmultiply. If all these operations could be integrated onto a single substrate considerable savings in power, weight and size could be realized. This section describes the design of a completely integrated 32-point CCD CZT IC which contains

- 4 64-stage CCD filters
- Clock logic and clock drivers
- Multiplying D/A converters (MDACs)
- ROM to store premultiply and postmultiply chirps.

In an effort to provide the greatest possible flexibility, the IC was partitioned into three main sections. The first and second sections, shown in Figure 9 and 10, contain two MDACs each. The internal ROM and the external ROM inputs are available to both MDAC sections. ROM select circuitry allows each MDAC to multiply the analog input signal by either an internal or an external signal. The internal ROM and the up/down counter used to address it are shown in Figure 9.

The main section of the IC is contained in section three, which is shown in Figure 11. This contains the CCD input amplifiers, which have provisions for blanking the inputs to the CCD; the 64-bit CCD transversal filters, which perform most of the computation required for the CZT; and the differential current integrators (DCI) used as the output amplifier for the CCD filters. Buffer amplifiers are provided for the outputs of the DCIs.

With this three-section scheme, any one section can be used independently with any other section or chip to implement several signal processing applications.

Integration of MDACs

The implementation of the CZT requires the input data to be premultiplied by a chirp waveform before filtering in a CCD. This multiplication can be easily implemented with MDACs based on a capacitive divider network (McCreary, Gray, 1975) approach. Each MDAC output is an analog signal whose value is proportional to the product of an analog input signal and a digital word. In addition to performing the pre- and postmultiplys needed for the CZT algorithm, the MDACs can also be used to perform multiplication by a reference signal for correlation applications. The area occupied by an 8-bit MDAC can be very small (0.6×0.9 mm^2), and the power required can be small if the load is a small CCD input capacitance.

ROM

The CCD CZT IC contains two MOS ROMs which supply the cosine and sine chirp signals to the MDACs. Since the chirp signal is symmetrical about the 16th point each ROM contains 17, 8-bit words and is addressed by a 17 stage up/down ring counter. A bit is coded as a "0" or a "1" by the presence or absence of a transistor, respectively.

CCD Filters

The CCD filters are 64-stage, surface channel filters implemented using the split electrode technique. (Buss, Collins, Bailey, Reeves, 1973) Like the other on-chip circuitry the CCD filters are N-channel silicon gate structures. To aid in testing the CCD filters separately, a pair of buffer amplifiers with differential outputs are used between the MDACs and the CCD inputs.

Differential Current Integrator

The function of the differential current integrator (DCI) is to integrate the clock line current differentially and provide an output signal. This differential current integration can be performed by integrating the CCD output signal onto feedback capacitors of an operational amplifier. A high gain MOS op-amp was designed for the DCI function. The amplifier has two stages of differential gain. Each stage is cascaded and uses depletion loads. The amplifier has internal compensation capacitors. Some of the important amplifier parameters are

- Open loop gain   1635
- Crossover frequency   17 MHz
- Closed loop gain   32
- Gain bandwidth product   550 MHz
- Phase margin   65°

The buffer amplifiers used as the inputs to the CCD filters have the same design as the DCI amplifiers except for the value of the compensation capacitors. The DCI amplifier is compensated for a higher closed loop gain.

Clock Generators

All control pulses needed for all the circuitry are generated on-chip. The IC requires only two input MOS level master clocks. The circuit configuration used to produce most of the clock pulses requires a pull up pulse and a non-overlapping pull down pulse. Each circuit uses a bootstraped approach to achieve fast rise times. (Typically 20 ns to 30 ns). Because the device sizes required for each generator depend on the capacitive load it must drive, each clock generator required a custom design.

Circuit Configurations

The CCD CZT IC is a general purpose IC that can make the CCD CZT practical for many system implementations. A few of the more general circuit configurations which utilize this IC are shown in Figures 12 through 15.

Figure 12 shows a system for obtaining the power density spectrum of a real input signal. This configuration has a 50% input/output duty cycle.

With a few external components, a complex input/output CZT can be performed with one chip as shown in Figure 13. The external components are necessary to disable the inputs when a postmultiply operation is needed for the true CZT. Because of this blanking, the system in Figure 13 has only a 50% input/output duty cycle. To obtain a 100% input/output duty cycle CZT, two chips are needed as shown in Figure 14. This configuration does not require the external circuitry of Figure 13.

The realization of a 16-point correlator using two CZT ICs is shown in Figure 15. This correlator convolves the real input signal with an impulse response $h(t)$. True correlation is obtained, since an on-chip pulse T_1 can blank the CCD input when required.

4.   APPLICATIONS AND CONCLUSIONS

For a given spectral analysis application to be considered as a candidate for CCD CZT implementation it must satisfy two criteria: (1) It must be of modest performance which lies within the CCD performance limitations and (2) It must be required in sufficiently high volume that low cost is a dominant design specification. These two criteria rule out a large class of applications. However, there have been identified, several applications of great importance which do satisfy both of the above criteria. These will be discussed in this section.

4.1   Video Bandwidth Reduction

Transform encoding of video images for the purpose of bandwidth reduction is a potentially important application of the CCD CZT. A hybrid transform system has been developed (Means, Whitehouse, Speiser, 1974) which performs a discrete cosine transform (DCT) in one dimension and differential pulse code modulation (DPCM) in the other dimension.

DFT and DCT on "typical" video images have resulted in variance compaction approaching that of optimum transforms, and both the DFT and DCT can be cost effectively implemented with the CCD CZT. The CCD CZT becomes particularly attractive in remote sensing applications such as RPV's where small size, light weight, and low power are essential in addition to low cost.

4.2   Speech Processing

Spectral analysis is one of the most important functions in speech processing, (Schafer, Rabiner, 1975) and speech processing requirements in terms of sample rate (10 kHz), delay time (40 ms) and dynamic range (40 dB) are well within the CCD capabilities outlined above.

There exist several algorithms for speech bandwidth reduction but one is particularly well suited to CCD CZT implementation. It is called homomorphic deconvolution (Oppenheim, Schafer, 1968) and operates upon the principle of the deconvolution of speech into pitch and vocal tract resonances. In Figure 16, a block diagram of such a system is shown. Speech (point A) is a quasi-periodic waveform with period equal to the pitch period (8 msec in this example corresponds to a pitch of 125 Hz). The log magnitude spectrum of speech (point B) consists of an envelope representing the vocal tract resonances which modulates all harmonics of the pitch. The cepstrum (point C) is the Fourier transform of the log magnitude spectrum. It has a peak at large time value (8 msec in this example) corresponding to the pitch period and a low time portion corresponding to the transform of the envelope of the log magnitude spectrum. By appropriately windowing the cepstrum, the pitch and vocal tract parameters can be determined and coded for data transmission. Homomorphic deconvolution is costly to implement digitally in real time because of the three sequential transforms which are required. However, the CCD CZT holds the promise of truly low cost implementations of this type of processing.

Figure 17 shows the operation of a cepstral unit in which the CZT blocks are implemented using the 500-point sliding CZT described in Section 3.1. The oscilloscope photograph on the right is a time expanded version of the one on the left. In each photograph, input speech (point A in Figure 16) is shown in the top trace, the log magnitude spectrum (point B in Figure 16) is shown in the center trace, and the cepstrum (point C in Figure 16) is shown in the bottom trace. Harmonics of the pitch frequency are clearly visible in the center trace and give rise to the impulse at the pitch period (about 5 msec) in the bottom trace.

4.3   Doppler Processing in MTI Radar

MTI (moving target indicator) radar operates upon the principle of detecting moving targets of small cross section in the presence of stationary background having much larger cross section. The doppler shift of the radar return is determined, and from this, the target velocity parallel to the radar line of sight, can be determined.

Radar returns are quasi-periodic, and the sliding CZT can be used. Typical transform lengths are 10 to 100 and typical pulse repetition frequency (PRF) is 1 kHz to 100 kHz. A doppler processor may be required to process thousands of DFT's in parallel, (one for each range gate) so reducing the cost of the DFT has a large cost impact on the overall system.

A doppler processor IC has been developed (Bailey, Hite, Eversole, McCray, 1977) which performs five 80-point CCD CZTs of the type shown in Figure 5. The IC contains all the integrated amplifiers and squaring circuitry required to obtain the power density spectrum in each range bin. A doppler processor for thousands of range bins can be implemented by cascading 5-bin IC's at a projected cost of approximately one-third that of an all digital processor designed using state-of-the-art digital hardware.

4.4   Conclusions

Other potential applications of the CCD CZT include processing for FLIR images, sonobuoy signal processing, and remote surveillance.

The CCD CZT is not expected to make the digital FFT obsolete in signal processing systems. However, for those spectral analysis applications which fulfill the twin requirements of modest performance and high volume, tremendous cost advantages can be gained using the CCD CZT. More applications will certainly emerge, but in the meantime, the potential cost impact in the application areas already identified guarantees the importance of the CCD CZT in spectral analysis signal processing systems of the future.

REFERENCES

Bailey, W.H., Hite, L.R., Eversole, W.L., and McCray, J.A., "Radar Video Processing with Charge Coupled Devices," to be presented at the Technical Seminar on Sensor Technology for Battlefield and Physical Security Application, 13-15 July 1977, Fort Belvoir, Virginia.

Brodersen, R.W., Hewes, C.R., and Buss, D.D., "A 500-Stage CCD Transversal Filter for Spectral Analysis," IEEE J. Solid-State Circuits, SC-11, pp 75-84, February 1976.

Buss, D.D., Collins, D.R., Bailey, W.H., and Reeves, C.R., "Transversal Filtering Using Charge Transfer Devices," IEEE J. Solid-State Circuits, SC-8, pp 134-146, April 1973.

Buss, D.D., Veenkant, R.L., Brodersen, R.W., and Hewes, C.R., "Comparison Between the CCD CZT and the Digital FFT," CCD '75 Proceedings, pp 267-281, San Diego, October 1975.

McCreary, J.L. and Gray, P.R., "All MOS Charge Redistribution Analog-to-Digital Conversion Techniques -Part I," IEEE J. Solid-State Circuits, SC-10, pp 371-379, December 1975.

Means, R.W., Whitehouse, H.J., and Speiser, J.M., "Television Encoding Using a Hybrid Discrete Cosine Transform and a Differential Pulse Code Modulator in Real Time," 1974 National Telecommunications Conference Record, San Diego, December 1974, pp 69-74.

Oppenheim, A.V., and Schafer, R.W., "Homomorphic Analysis of Speech," IEEE Trans. Audio. Electroacoustic, Vol AU-16, pp 221-226, June 1968.

Rabiner, L.R., and Gold, B., Theory and Application of Digital Signal Processing, Prentice-Hall, 1975, Ch. 6.

Rabiner, L.R., Schafer, R.W., and Rader, C.M., "The Chirp Z-Transform Algorithm," IEEE Trans. on Audio. and Electroacoustics, AU-17, pp. 86-82, June 1969.

Rader, C.M., "Discrete Fourier Transforms When the Number of Data Samples is Prime," Proc. IEEE 56, pp 1107-1108, June 1968.

Schafer, R.W. and Rabiner, L.R., "Digital Representations of Speech Signals," Proc. of the IEEE 63, pp 662-677, April 1975.

Welch, P.D., "A Fixed-Point Fast Fourier Transform Error Analysis," IEEE Trans. Audio Electroacoustic, AU-17, pp 151-157, June 1969.

Whitehouse, H.J., Speiser, J.M. and Means, R.W., "High Speed Serial Access Linear Transform Implementations," presented at the All Applications Digital Computer Symposium, Orlando, Florida, 23-25 January 1973, NUC TN 1026.

ACKNOWLEDGEMENT

TABLE 1

CCD CZT PERFORMANCE LIMITATION

Parameter	Limitation	Value
Transform Length	Imperfect CTE Chip Size	1000
Time Record	Leakage Current	1 sec
Speed	Analog Circuitry	5 MHz
Accuracy	Premultiply Quanti- zation; Weighting Coefficient Error	13 - Bit Equivalent FFT
Resolution	Imperfect CTE	5% Degradation
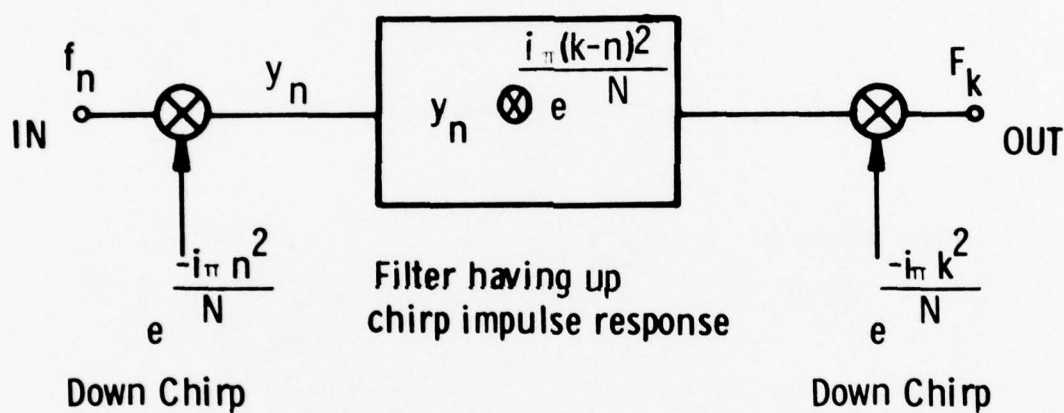
Fig.1   Schematic of the CZT algorithm
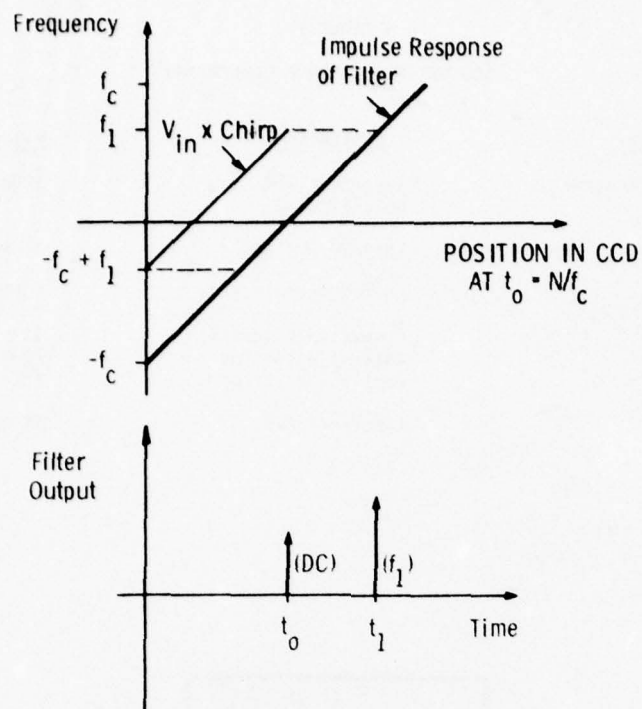
Fig.2 Interpretation of the CZT in terms of chirp input waveforms in chirp filters
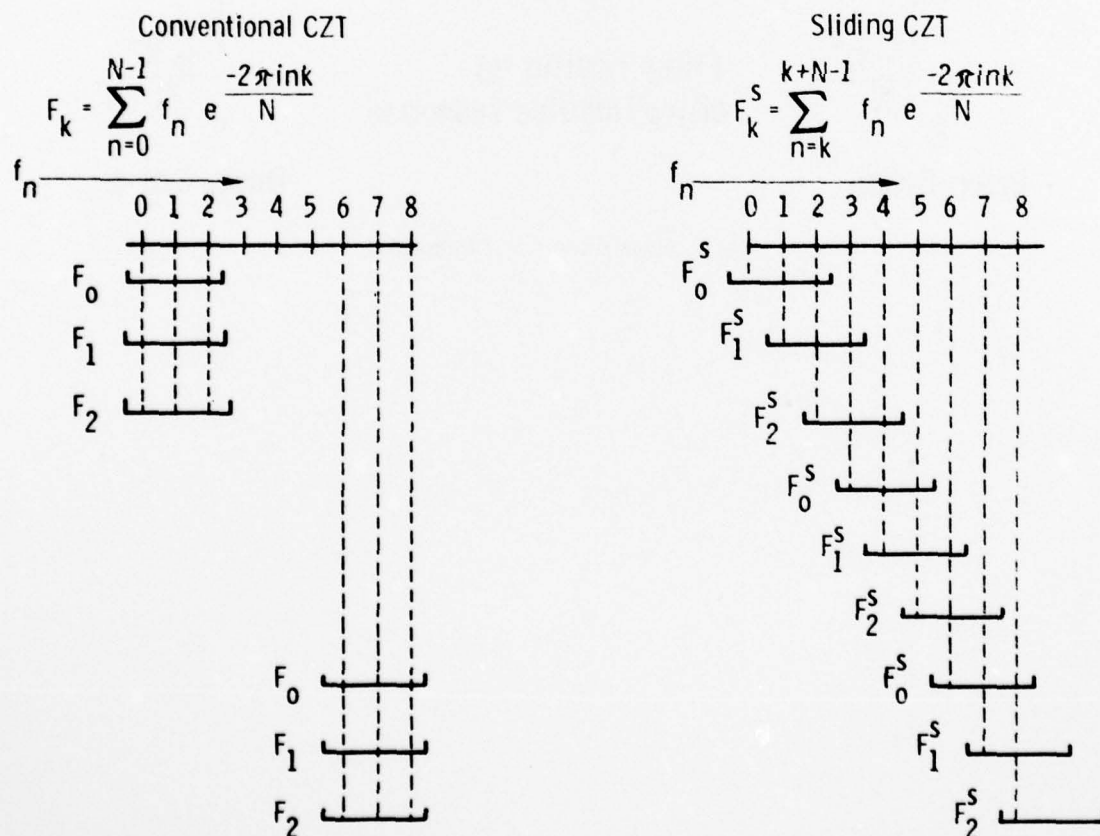
Fig.3 Comparison between the conventional CZT and the sliding CZT for the simple example of a 3-point transform
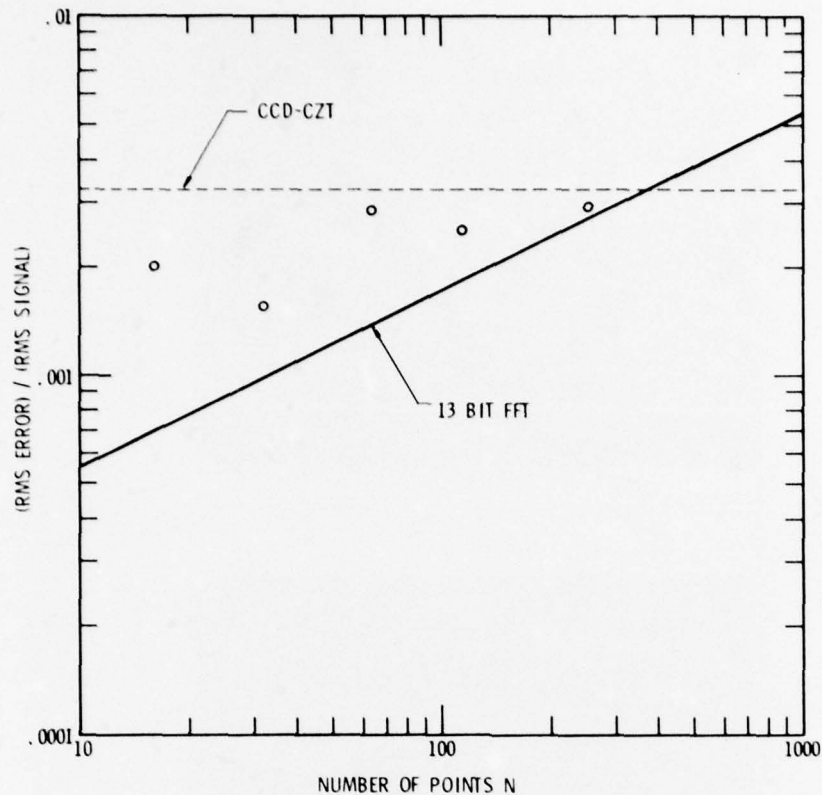
Fig.4   Error comparison between a CCD CZT and a digital FFT implemented using 13 bits plus sign.
The CCD CZT is limited by the quantization of multiply chirps to 7 bits plus sign. The points represent computer
simulation for the CCD CZT
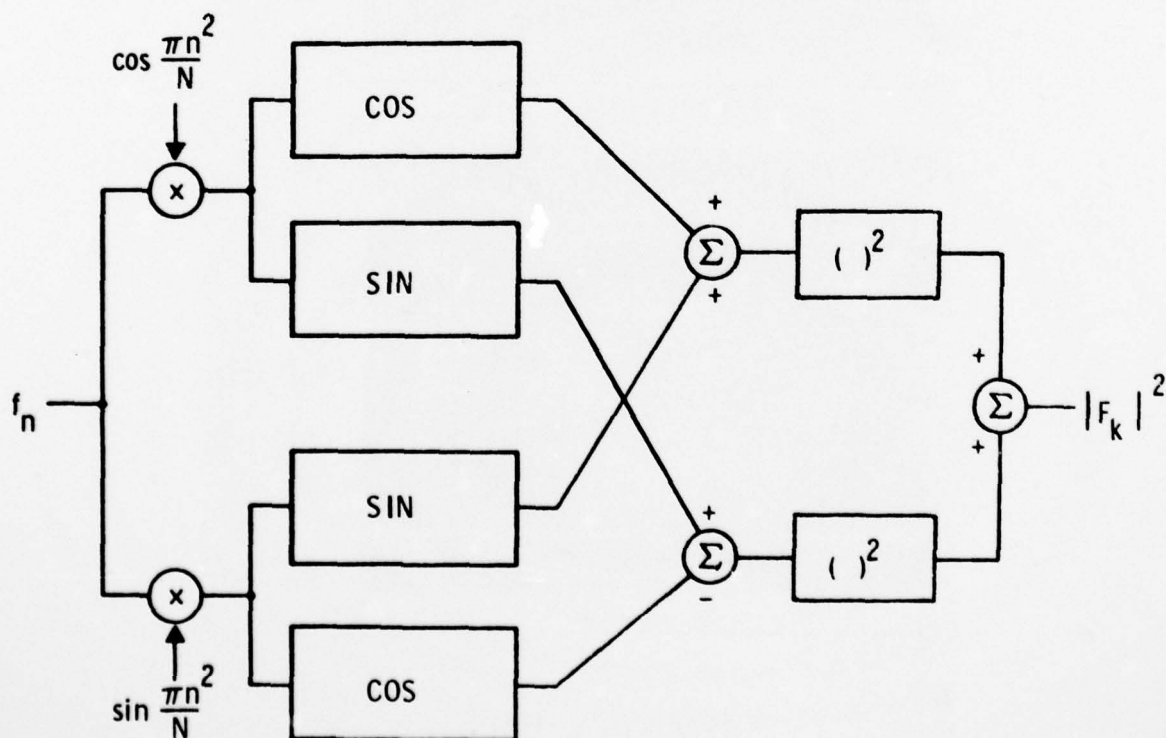
Fig.5   Block diagram of the complex arithmetic of the CZT algorithm for computing power density spectrum. The
CCD transversal filters are designated by COS and SIN.

INPUT
200 Hz

2.5 kHz

4.5 kHz

FREQUENCY (kHz)
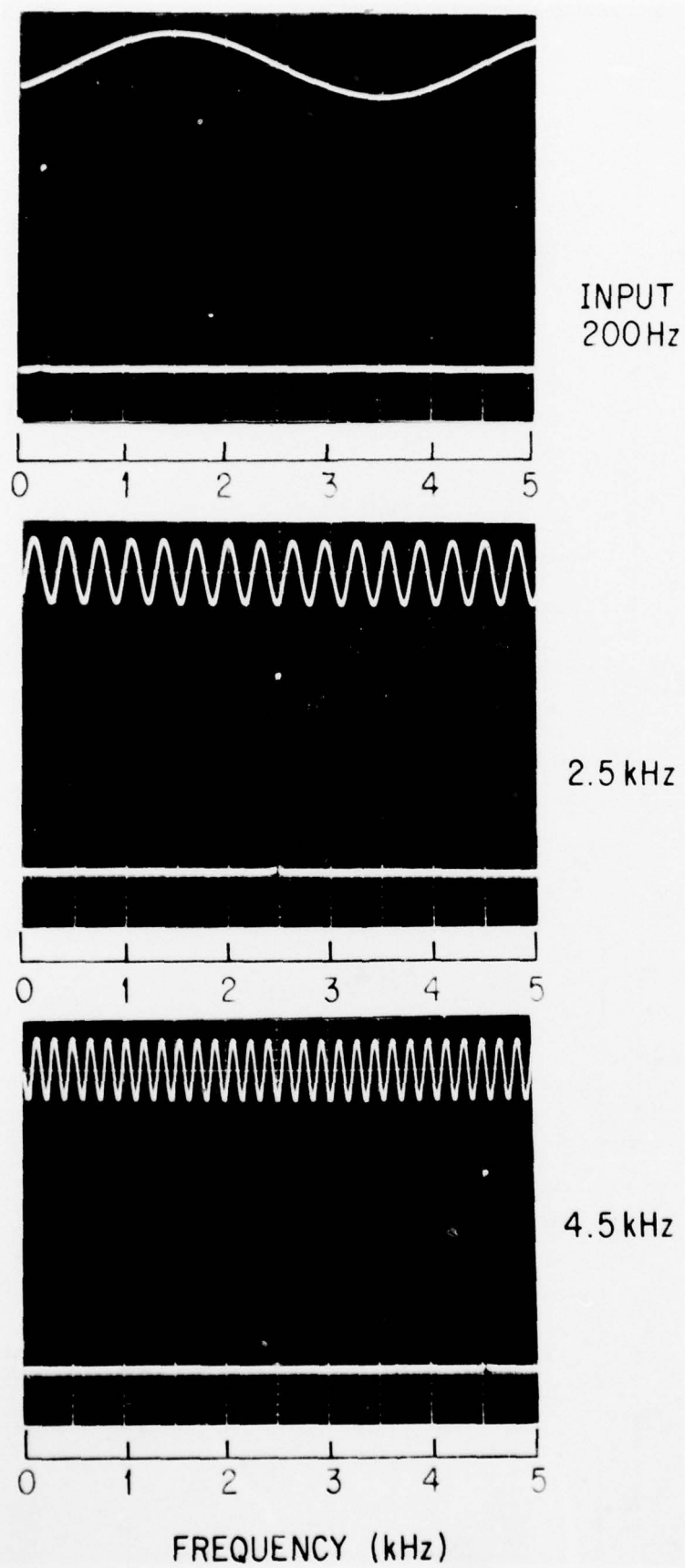
Fig.6   Power density spectra for three sinewaves obtained using the 500-point sliding CCD CZT at a sample rate of
20 kHz.

(a) INPUT 200Hz

(b) OUTPUT SPECTRUM x1

f_0 3rd 5th 7th 9th

(c) x10

f_0 3rd 5th 7th 9th

(d) x100

11th 13th 15th 17th 19th
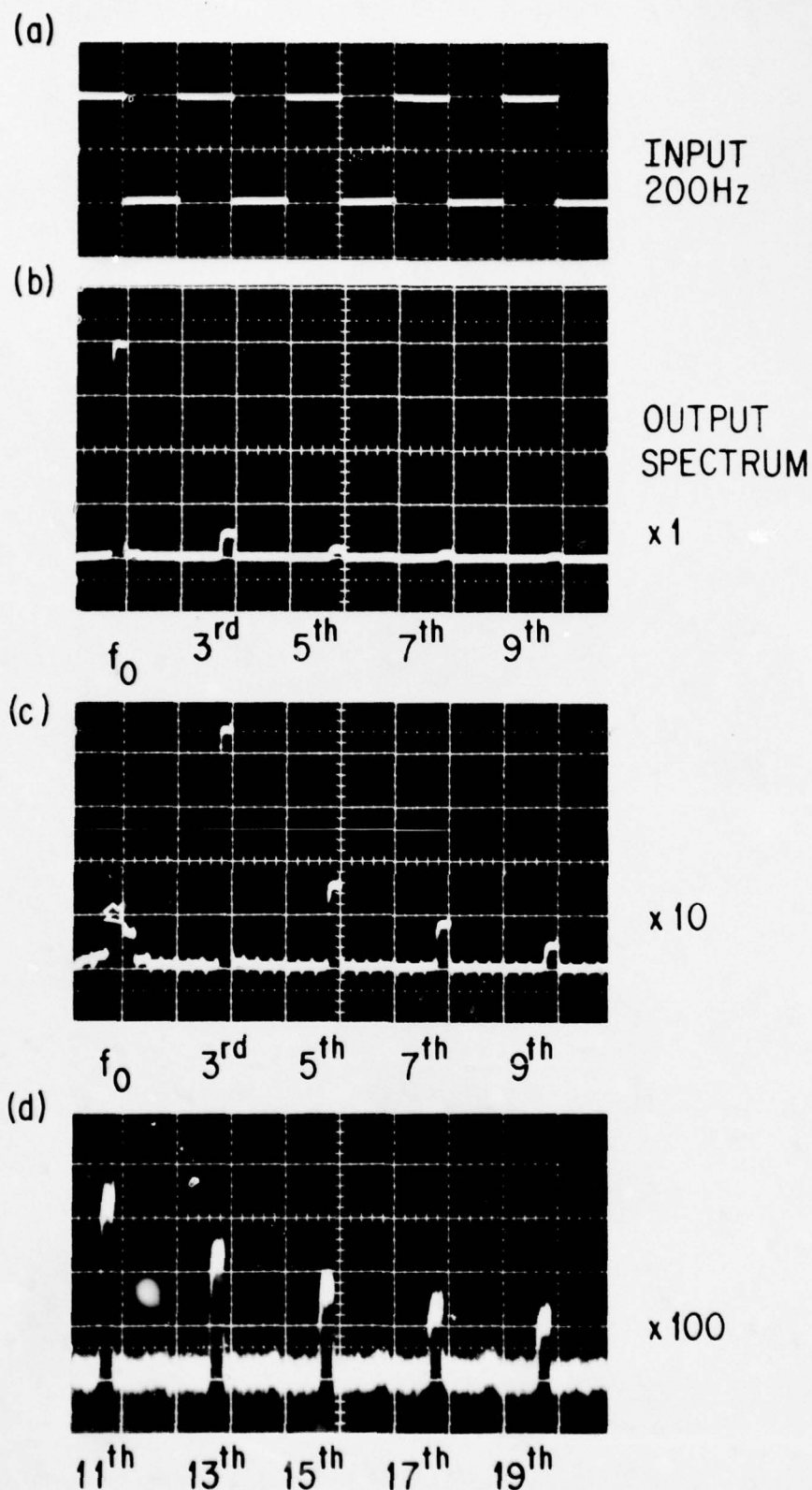
Fig. 7   500-point power spectrum of a 200 Hz square wave at a sample rate of 25 kHz. (a) Input waveform; (b) power spectrum output showing the first nine harmonics; (c) the output amplified by 10, which shows a small trailing pulse due to transfer inefficiency (pointed out by the arrow); and (d) the 11th through 19th harmonics amplified by 100.
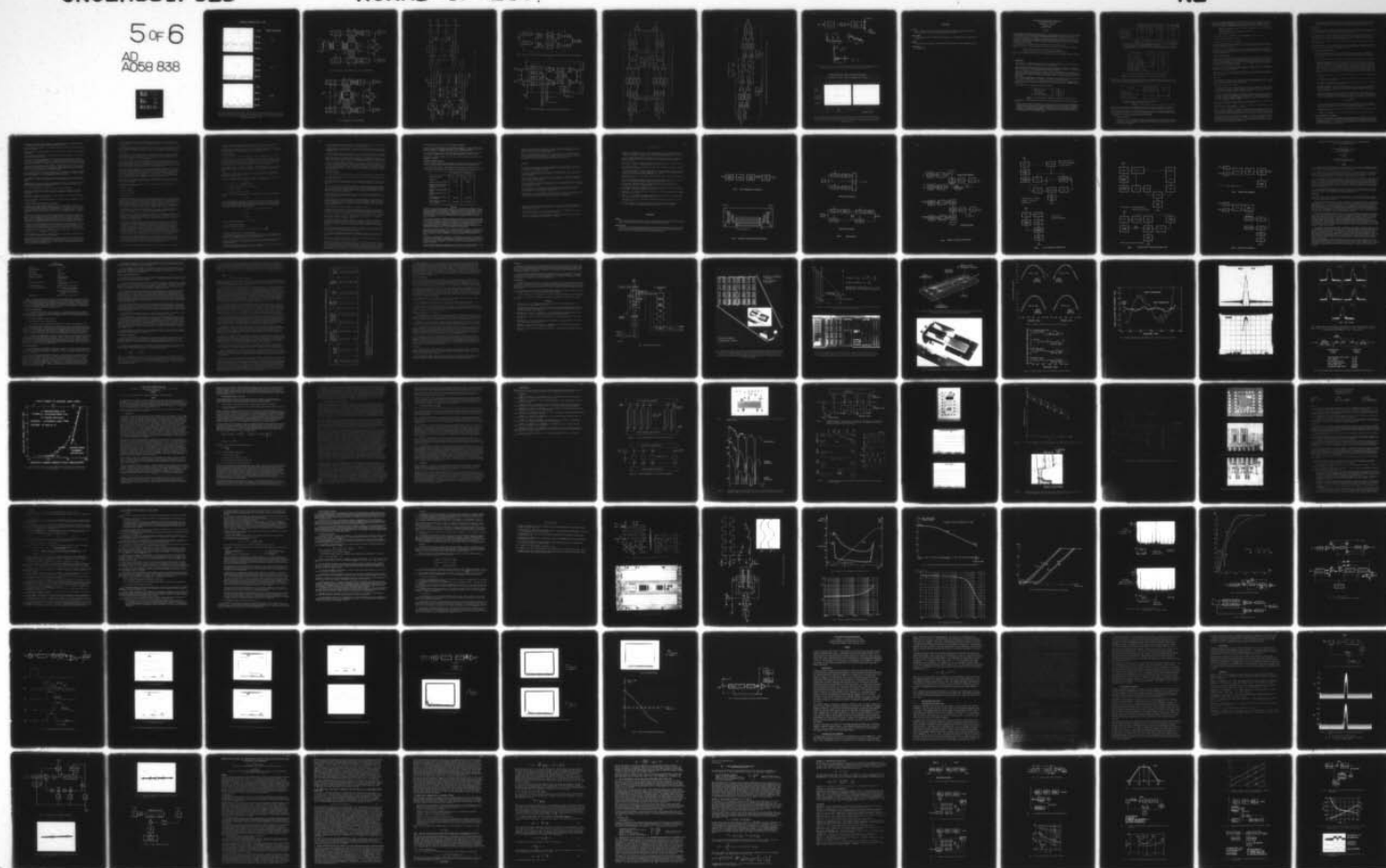
DYNAMIC RANGE OF CCD - CZT
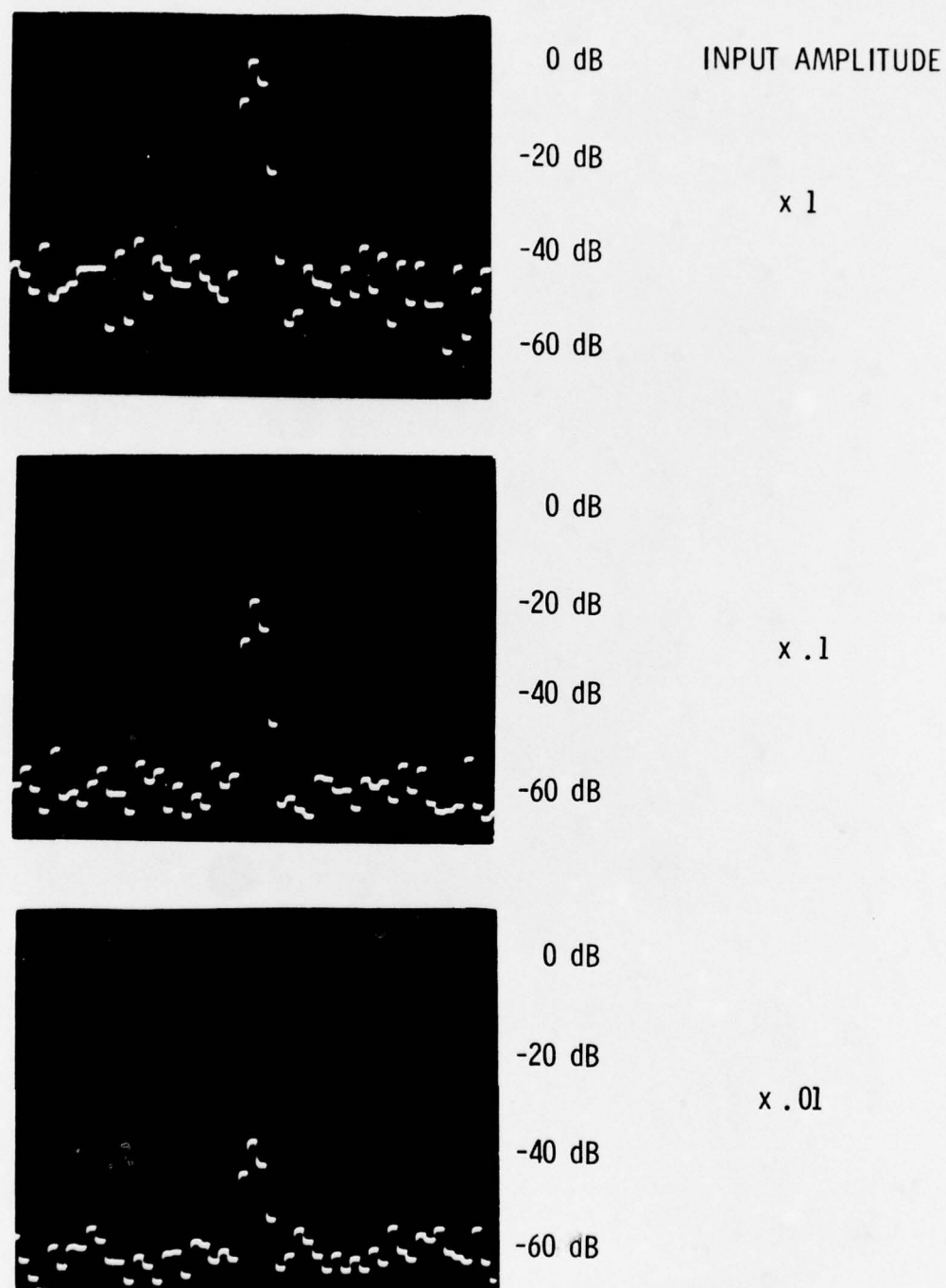
Fig.8 Output of the windowed CZT unit when a single frequency sinusoid is applied to the input. The top photograph shows — 40 dB sidelobes when the maximum signal amplitude is applied. The center and bottom photographs show the output when the input is attenuated by 20 dB and 40 dB, respectively. The noise level in the bottom photograph is approximately — 60 dB.

Fig.9 Block diagram of internal ROM, up/down counter, and dual MDACs.

Fig.10 Block diagram of second dual MDAC.
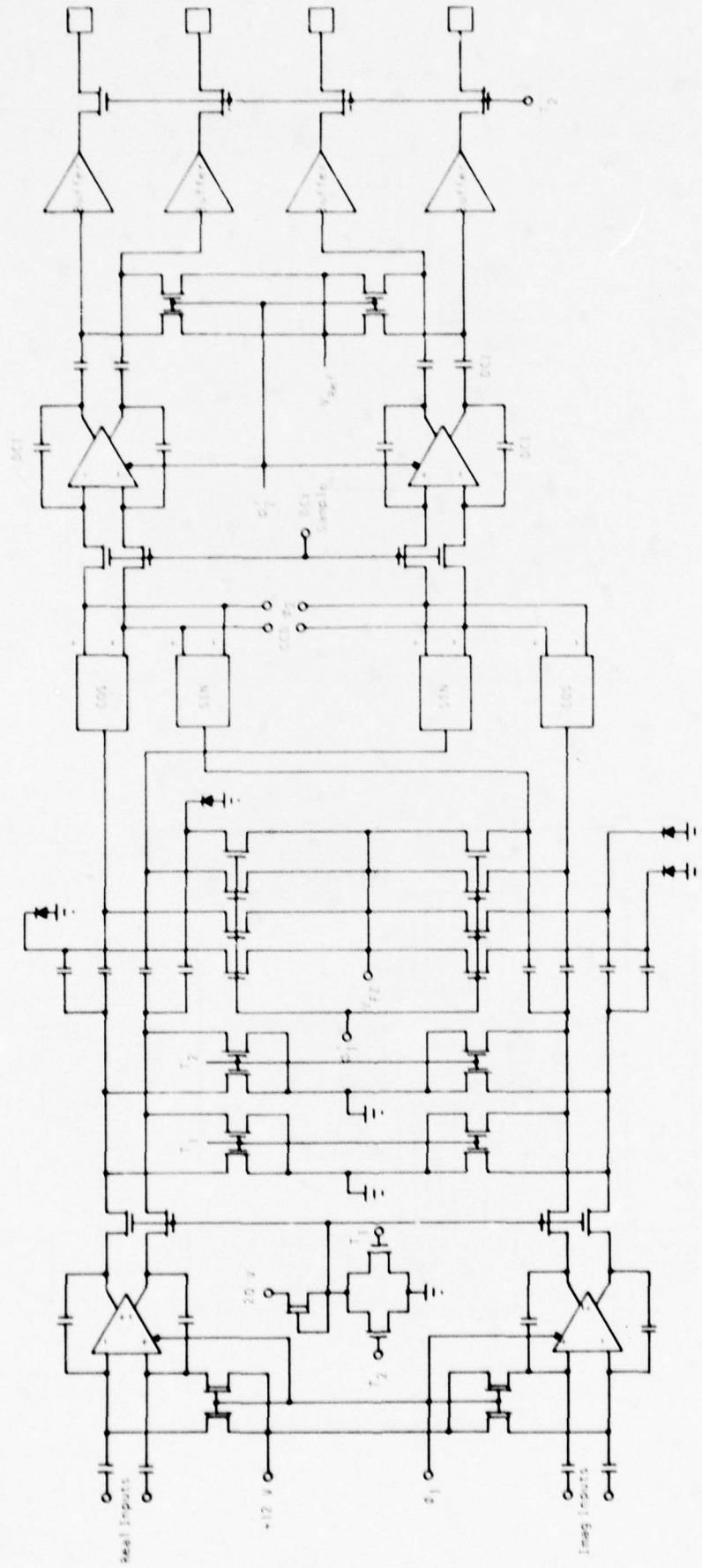
Fig 11 Block diagram of CCD input amplifiers, CCDs, DC1, and output buffers.
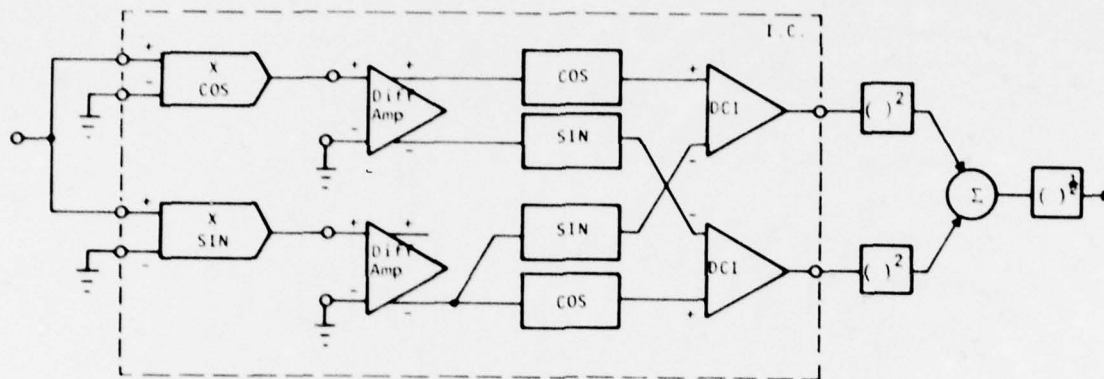
Fig.12   A system for obtaining power density spectrum of a real input using one CZT IC (50% input/output duty cycle).
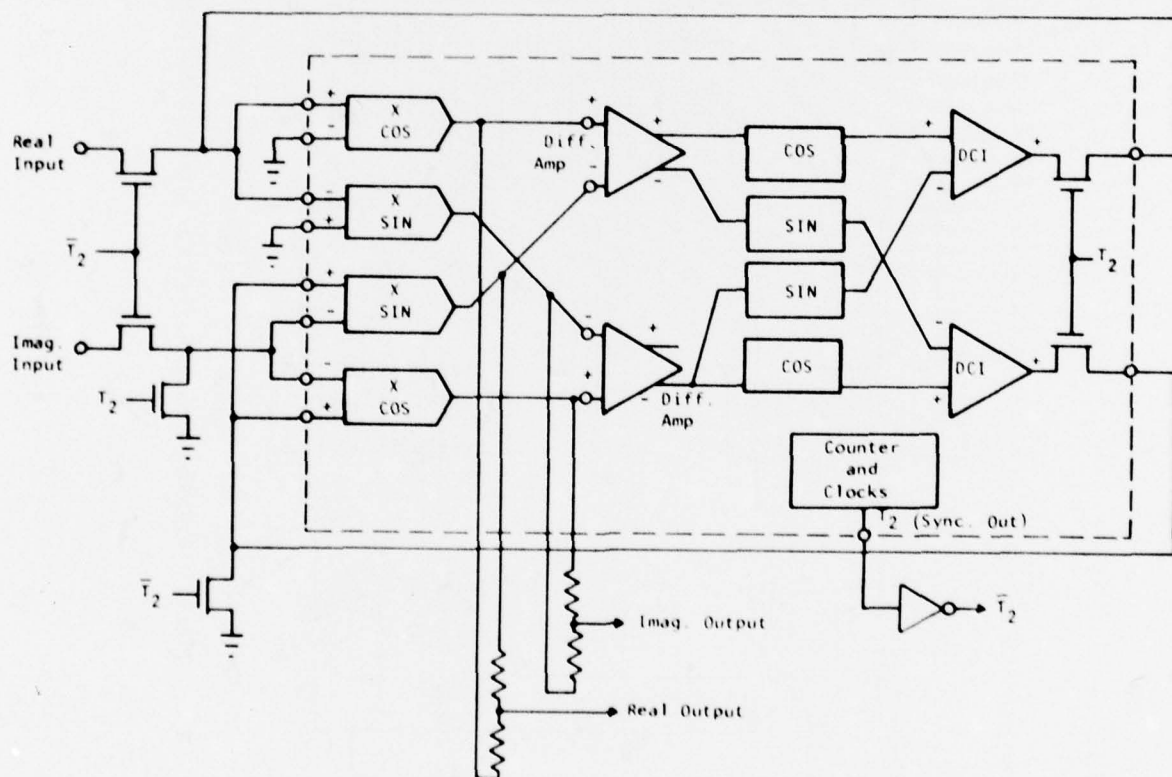
Fig.13   Realization of complex I/O CZT with one chip (50% I/O duty cycle).
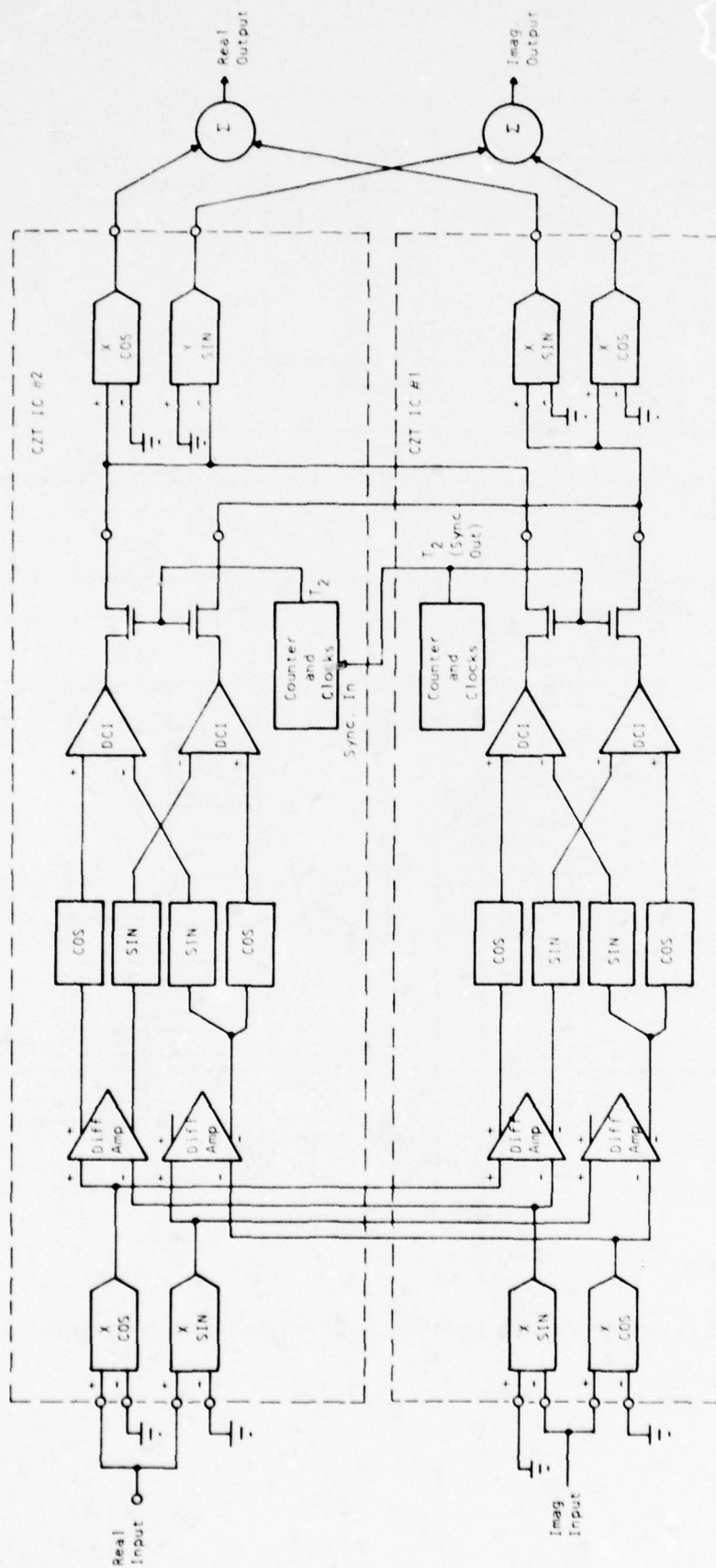
Fig.14   Realization of complex I/O CZT using two CZT ICs (100% I/0 duty cycle).
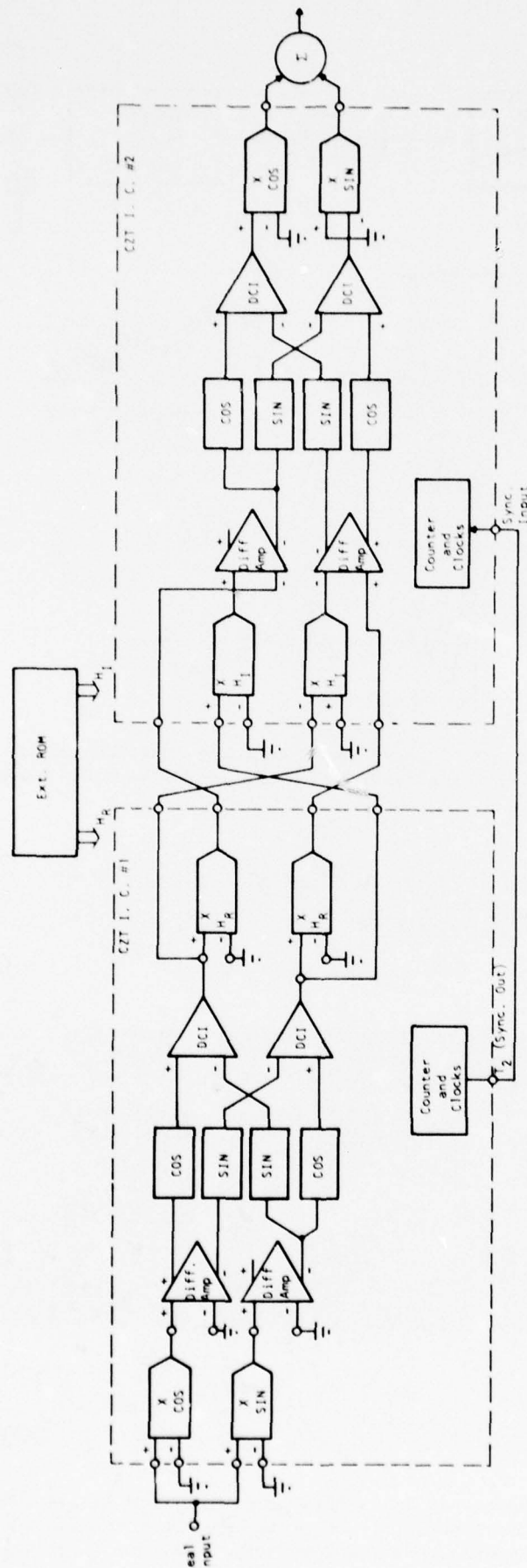
Fig.15   Realization of correlator using two CZT ICs (Real 25% duty cycle input, real 50% duty cycle output).
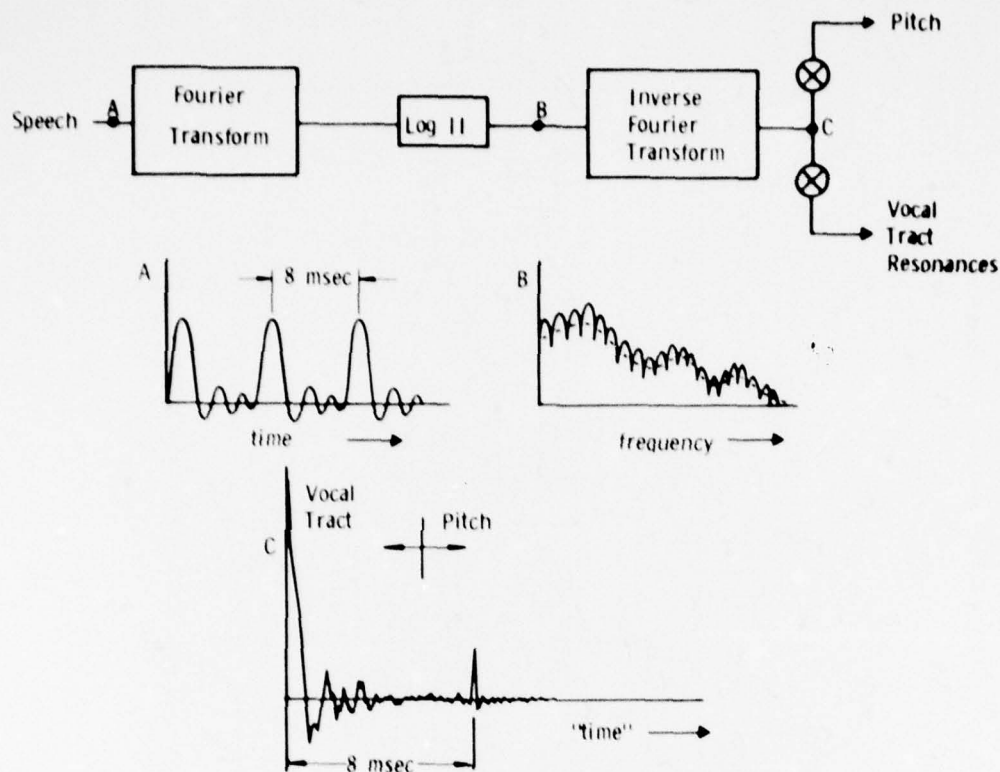
Fig.16 Block diagram of a system to perform homomorphic deconvolution of speech for bandwidth reduction. This type of speech processing is particularly amenable to implementation with the CCD CZT.

SPEECH PROCESSING WITH CHARGE-COUPLED DEVICES

Homomorphic Processor: Cepstrum Generation-Pitch Detector
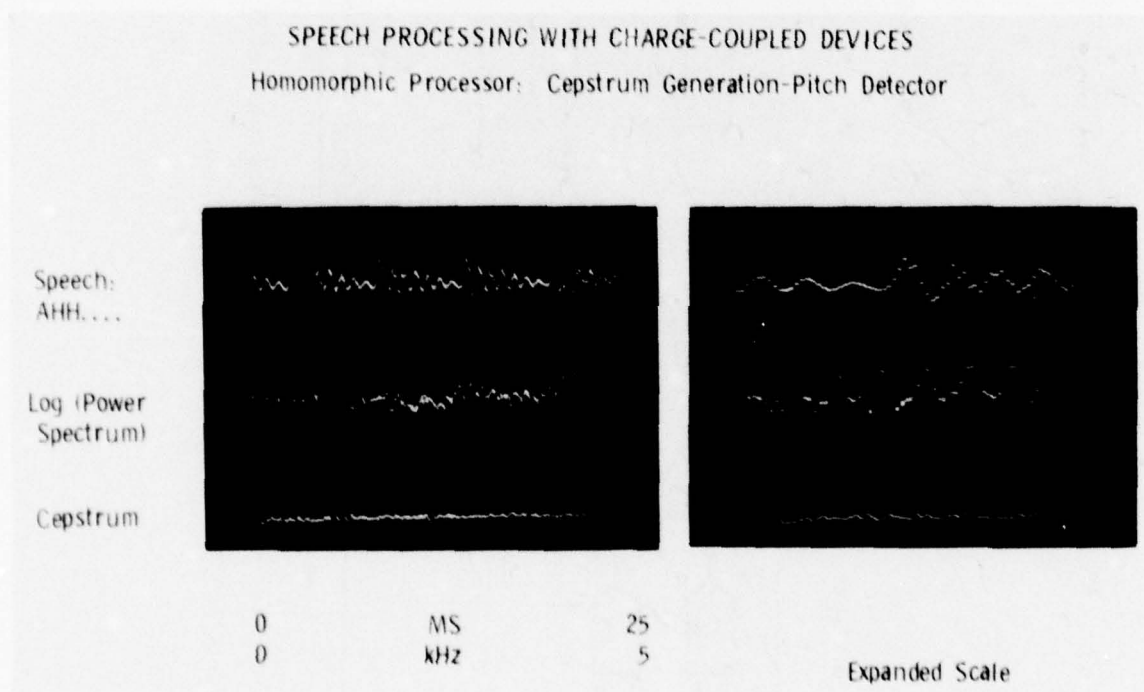
Fig.17 Operation of the homomorphic deconvolution system of Figure 16 using the 500 point CZT discussed in Section 3.1. Speech (point A) is shown in the top trace, the log magnitude spectrum of speech (point B) is shown in the center trace, and the cepstrum (point C) is shown in the bottom trace. The photograph on the right has an expanded time base.

DISCUSSION

Tournois

You are speaking about sonar and spectrum analysis applications. What is the minimum clock rate you can obtain on your device, or what is the minimum bandpass you can analyze?

Author's Reply

For the 500 stage we operate at 500 Hz, which corresponds to 1 sec storage time which corresponds to 1 Hz resolution.

Roberts

For doppler radar processing did you find a good uniformity in response for from one range bin to the next?

Author's Reply

It is fairly uniform.

APPLICATIONS D'UN CONVOLUTEUR PIEZOELECTRIQUE
AU TRAITEMENT DU SIGNAL RADAR

P. ANTHOUARD - T. BEAUVAIS

THOMSON-CSF
Malakoff, France

RESUME

Plusieurs familles de dispositifs à ondes de surface ont été étudiées au cours des dix dernières années. Ces dispositifs sont essentiellement des composants analogiques haute fréquence mais ils peuvent être avantageusement associés à des systèmes de traitement numérique pour certaines applications.

L'utilisation d'un convoluteur piézoélectrique appliqué au traitement de corrélation de signaux échantillonnés et quantifiés en est un exemple. Ce convoluteur permet de très grandes vitesses de calcul difficiles à obtenir par des techniques entièrement numériques. Néanmoins, il nécessite des interfaces analogique-numérique rapides.

La première partie de l'article décrit le convoluteur et les interfaces associées.

La seconde partie montre quelques exemples d'association du convoluteur à des chaînes de traitement radar :

- mesure de temps d'arrivée, tri et identification de signaux,

- traitement du signal issu d'un radar à antenne latérale synthétique.

Dans ce dernier exemple, l'accent est mis sur la comparaison entre les traitements acoustique et numérique.

1 - INTRODUCTION

L'exploitation des signaux radars fait appel à des traitements complexes qui nécessitent d'être de plus en plus précis, rapides et adaptés à des situations variées.

Dans ce domaine, les fonctions de corrélation et convolution jouent un rôle important, longtemps rempli, à l'exception du filtrage adapté spécifique, par les techniques numériques, qui offrent l'avantage de la précision et de la flexibilité pour une complexité moyenne.

Devant l'accroissement des contraintes de vitesse de calcul, d'encombrement et de consommation, en particulier dans les systèmes aéroportés, l'intérêt s'est porté vers d'autres technologies telles que les composants acoustiques spécialisés, particulièrement adaptés au traitement rapide des signaux analogiques.

Dans le domaine du traitement de corrélation et convolution, les systèmes utilisables sont les dispositifs à effets paramétriques et les dispositifs utilisant des filtres dispersifs.

Ces derniers ont été décrits avec leurs applications dans différentes publications [1],[2],[3] ; la fonction de base d'un corrélateur (ou convoluteur) à filtres dispersifs est la transformée de Fourier ("Chirp Z").

Les performances de ces systèmes sont résumées au tableau 1 :

Nombre de points	< 5000
Bande passante	< 600 MHz
Durée de traitement	< 100 ms

TABLEAU 1 : Caractéristiques des systèmes de traitement à filtres dispersifs

Les dispositifs à effets paramétriques constituent une autre approche du traitement de corrélation, ce sont :

1) les corrélateurs à mémoire actuellement à l'étude dans plusieurs laboratoires [4],[5],[6],[10]. Ce type de composant est constitué d'un substrat piézoélectrique au voisinage duquel est disposé un réseau de diodes intégrées dans un matériau semiconducteur permettant le stockage des signaux appliqués. Il combine les fonctions mémoire et corrélation, et son étude devrait déboucher sur un composant utilisable dans quelques années. Le tableau 2 montre les performances obtenues sur de tels dispositifs et leurs limites réalistes.

	Corrélateur à diodes p-n[3]		Corrélateur à diodes Schottky[6][4]	
	Démontré	limites	Démontré	limites
Bande passante	20 MHz	100 MHz	30 MHz	100 MHz
Durée de signaux	6 µs	10 µs	10 µs	-
Temps de mémoire	5 s	10 s	10 ms	100 ms
Pertes en filtre à corrélation	- 50 dB	-	- 50 dB	-

TABLEAU 2 : Caractéristiques techniques des corrélateurs à mémoire

2) Les convoluteurs piézoélectriques largement développés dans plusieurs pays [7], [8], [9], [10] et qui constituent un composant utilisable actuellement.
Les applications décrites ci-après utilisent l'un d'eux, développé dans les laboratoires de THOMSON-CSF ; le tableau 3 montre ses performances actuelles et ses limites futures.

Caractéristiques	actuelles	futures
Temps de traitement	12 µs	20 µs
Bande à 3 dB à l'entrée	50 MHz	100 MHz
Uniformité de traitement	± 0,2 dB	± 0,2 dB
Fréquence de fonctionnement	156 MHz	300 MHz
Facteur de bilinéarité*	- 72 dBm	- 72 dBm
Suppression de la convolution de trajet double	40 dB	60 dB
Puissance maximale à l'entrée	1 W	1 W

* Pic de convolution pour signaux d'entrée de niveau électrique 0 dBm

TABLEAU 3 : Caractéristiques techniques du convoluteur piézoélectrique.

Les utilisations décrites du convoluteur portent sur des traitements d'identification, de mesure de temps d'arrivée et de traitement des signaux issus d'un radar à antenne latérale synthétique. Le tableau 4 en résume les principales conditions d'utilisation.

Application	Signaux à traiter		Remarques
	Durée Θ	Bande β	
Identification	< T	< B	
Mesure de temps d'arrivée entre - τ et + τ	> T + 2τ	< B	Précision Δτ<1/B
Récepteur à corrélation	< T	< B	

T : temps de propagation sous l'électrode de convolution
B : bande passante du convoluteur à l'entrée

TABLEAU 4 : Conditions d'utilisation du convoluteur piézoélectrique

Ces applications supposent un fonctionnement synchrone du convoluteur et nécessitent une adaptation temporelle entre les signaux issus du récepteur de base et les signaux appliqués au convoluteur.

Cette adaptation est obtenue par la mise en oeuvre d'interfaces : convertisseurs analogique-numérique, numérique-analogique et mémoire, qui rendent le convoluteur adaptable à une gamme de signaux étendue.

2 - DESCRIPTION DU CONVOLUTEUR PIEZOELECTRIQUE ET DES INTERFACES

2.1. Chaîne de traitement

Une chaîne de traitement typique de signaux radars, bâtie autour du convoluteur est représentée à la figure 1. C'est une combinaison d'un traitement numérique et analogique permettant de réaliser une réduction importante du volume et du coût des opérateurs par rapport à un traitement entièrement numérique.

La partie de traitement numérique qui précède sert d'interface entre les signaux issus du récepteur à fréquence intermédiaire et le convoluteur. Il a essentiellement pour but d'adapter les caractéristiques des signaux reçus à celle du convoluteur. C'est en général une mise en mémoire et une organisation spéciale de la lecture pour réaliser les fonctions suivantes :

- compression du temps pour adapter le spectre des signaux reçus à la bande passante du composant acoustique,

- inversion de temps pour les traitements de corrélation,

- multiplexage du convoluteur.

La complexité de ce sous-ensemble est liée à deux paramètres principaux qui sont :

1) le nombre d'éléments binaires à traiter,

2) la fréquence de codage de cette information.

La partie de traitement analogique comporte un convertisseur numérique-analogique qui reçoit les signaux numérisés à fréquence porteuse nulle et les délivre après codage à la fréquence porteuse de travail du convoluteur.

La difficulté dans cette chaîne de traitement est essentiellement posée par la fréquence de codage élevée, pour une dynamique donnée, aussi bien pour les codeurs que pour les mémoires associées. La dynamique du traitement est limitée actuellement à 40 dB environ pour une période d'échantillonnage de 10 ns.

2.2. Convoluteur piézoélectrique

Le convoluteur est constitué d'un substrat piézoélectrique poli (figure 2) sur lequel sont disposés deux transducteurs qui permettent de générer deux ondes acoustiques de surface à la fréquence ω qui se propagent en direction opposée et dont les modulations représentent respectivement les signaux à convoler $f(t)$ et $g(t)$. Des coupleurs à bandes métalliques compriment les faisceaux acoustiques dans un rapport 10 à 20 ; ceci augmente la densité de puissance acoustique et donc le rendement du système.

Du fait de la non-linéarité du matériau piézoélectrique, les deux ondes ainsi comprimées sont multipliées, et sur le triplaque de sortie un signal à la fréquence 2ω qui est l'intégrale du produit de f et g peut être détecté.

Dans le cas où f et g sont synchrones et de durée au plus égale au temps T de propagation sous l'électrode centrale, le dispositif génère le produit de convolution f • g comprimé dans le temps d'un facteur 2.

2.3. Interfaces

L'utilisation du convoluteur piézoélectrique à partir de signaux numérisés nécessite en premier lieu la modulation d'une porteuse à la fréquence centrale du dispositif acoustique (156 MHz) par les informations échantillonnées et quantifiées.

Deux types de procédés sont utilisables :

- le premier n'est cité que pour mémoire parce que limitatif en dynamique et bande passante, il ne permet pas d'utiliser toutes les capacités du convoluteur pour des traitements élaborés. Il consiste à moduler la fréquence porteuse en amplitude, par un signal analogique obtenu en sortie d'un convertisseur numérique-analogique.

- le second fait appel à la modulation d'une fréquence porteuse directement par les informations binaires. Deux types de modulateur sont applicables, qui peuvent être adaptés au traitement : modulation d'amplitude ou modulation de phase ; une configuration mixte, ou l'utilisation de deux modulateurs combinés en quadrature, peuvent être également adoptés pour les traitements vectoriels. (figure 3)

Le principe du modulateur consiste à séparer en N voies commutables, pour un codage à N éléments binaires, le signal de référence issu d'un oscillateur local.

Chaque voie comporte un atténuateur 2^i ($i = 0$ à $N - 1$) ou une cellule de déphasage $\Pi/2^i$ suivant qu'est effectué un codage d'amplitude ou de phase.

Parmi les types de commutateurs utilisables : commutateurs logiques, commutateurs à diodes Schottky, modulateurs équilibrés, commutateurs à diodes PIN, les commutateurs en logique sont les plus intéressants par leur encombrement réduit et leur simplicité de mise en oeuvre pour les applications à dynamique moyenne (40 dB) qui sont celles du convoluteur piézoélectrique.

Directement associées aux modulateurs dans la plupart des utilisations, il faut mentionner les mémoires numériques. Leur rapidité est compatible avec celle des commutateurs : des mémoires ECL de 1 K bits à temps d'accès de 20 ns sont actuellement disponibles en un seul circuit intégré ; leur utilisation combinée à celles d'horloges rapides permet d'atteindre des cadences d'échantillonnage inférieures à 10 ns.

Enfin, comme derniers éléments d'interface, il faut citer les circuits d'amplification néces-
saires à l'utilisation du convoluteur, et qui en compensent les pertes. Les valeurs typiques
à retenir sont des gains d'environ 40 dB et des puissances de sortie voisines de 0,5 W.

3 - APPLICATIONS

Les exemples de traitement de corrélation exposés ci-après, utilisant le convoluteur piézoélectrique,
sont tirés du domaine des traitements des signaux radars. Ils sont comparés en termes de faisabilité,
de rapidité et de précision, à des systèmes numériques équivalents.

D'une façon générale ces systèmes numériques équivalents utilisent la corrélation discrète, de struc-
ture analogue à celle des filtres transversaux, la corrélation rapide basée sur l'utilisation de
l'algorithme de la FFT n'étant en général intéressante que pour des traitements à grand nombre de
points qui restent du domaine des calculateurs spécialisés.

Ces exemples montrent l'intérêt du convoluteur piézoélectrique lorsque la vitesse de calcul est l'é-
lément essentiel.

A titre de comparaison les circuits numériques les plus rapides permettent d'effectuer le calcul d'un
point d'une fonction de corrélation dans le même ordre de grandeur de temps où toute la fonction est
reconstituée par le convoluteur (12 microsecondes).

3.1. Mesure de temps par corrélation

Une méthode de mesure de retard entre deux signaux à spectre limité, issus d'une même source
et ayant transité par deux voies différentes, peut-être effectuée en traçant la fonction de
corrélation appliquée à la modulation de ces signaux et en évaluant son centre par rapport à
une origine.

Les deux systèmes de traitement, numérique ou acoustique sont utilisables à cet effet (figure 4).

. corrélateur numérique : dans cette application la corrélation est effectuée non pas sur
l'amplitude des signaux, mais sur leur modulation de phase, ce qui a pour effet de s'affran-
chir de la dynamique de ces signaux et de simplifier les circuits de calcul. Pour améliorer
la précision, le système utilise un interpolateur qui reconstitue la fonction à partir d'un
nombre limité de points : dans une première période les échantillons de phase codés des deux
voies sont mémorisés sur une durée limitée ; dans la période qui suit, la fonction de corré-
lation est tracée en faisant glisser les échantillons d'une voie par rapport à l'autre. Elle
passe ensuite dans l'interpolateur pour l'évaluation précise du pic qui fournit la mesure du
retard.

La précision de mesure par ce procédé est une faible fraction de 1/B, où B et la bande passan-
te des signaux.

La vitesse de calcul, en logique ECL, d'un point de la fonction est voisin de 50 ns x N, où N
est le nombre de points mémoire.

Le volume et la consommation de l'opérateur de calcul sont d'environ 1 litre et 15 watts pour
une quantification de phase à 4 éléments binaires.

. corrélateur acoustique : suivant les mêmes principes, le convoluteur piézoélectrique avec les
interfaces associées est utilisable. Les mémoires numériques sont conservées pour assurer
la mise à l'échelle des signaux pour le traitement et l'inversion de temps à la lecture des
informations mémorisées.

La précision de mesure est un peu moindre que précédemment du fait des différents parasites
qui peuvent être introduits par le composant acoustique ; elle peut aussi être limitée par la
vitesse d'échantillonnage des circuits de mesure, compte tenu de la mise à l'échelle en temps
effectuée à la mémorisation. D'une façon générale cette précision de mesure est de la forme :

$$\sigma \approx \frac{1}{p\sqrt{2}} \sqrt{\frac{1}{S/B} + s^2}$$

où : p est la pente de la fonction de corrélation au seuil de mesure ; S/B est le rapport
signal à bruit thermique ; s^2 est une variance due aux autres erreurs que le bruit thermique
(quantification et ondulations en sortie du convoluteur dues aux parasites).

Le temps de calcul de la fonction de corrélation complète est fourni par le domaine de retard
du convoluteur (12 µs).

Le volume et la consommation de l'opérateur de calcul sont d'environ 1 litre et 5 watts.

3.2. Tri et identification de signaux

Une application du traitement de corrélation aux performances limitées jusqu'ici par les faibles
cadences de calcul des circuits numériques existants trouve son essor grâce aux convoluteurs-
corrélateurs acoustiques : il s'agit du tri des signaux radars par corrélation avec une réplique
préalablement mémorisée, ou l'identification d'un signal radar par corrélation avec un certain
nombre de répliques tenues en bibliothèque.

L'intérêt du convoluteur piézoélectrique est alors essentiel lorsque la cadence d'apparition des signaux est rapide, ou quand le nombre de répliques est élevé.

Tri : Un système de traitement de tri de signaux, utilisant le convoluteur piézoélectrique est présenté à la figure 5.

Les signaux sont reçus séquentiellement en sortie d'un récepteur et comparés à la réplique issue du même canal.

Une conversion analogique-numérique est rendue nécessaire pour la mémorisation de la réplique après normalisation des signaux.

Des circuits de traitement complémentaires à l'ensemble convoluteur-interfaces, décident de l'identité ou non des signaux corrélés. Ces circuits complémentaires permettent de calculer un coefficient de corrélation normalisé et de le comparer à un seuil voisin de 1 pour tenir compte des fluctuations fournies par l'échantillonnage, la quantification et le convoluteur lui-même.

Le domaine des signaux analysables est limité par l'échantillonnage pour les signaux à spectre large et par la précision de discrimination pour les signaux à spectre étroit. Ces limitations apportées par la technologie des circuits extérieurs interviennent actuellement simultanément avec les limitations apportées par le convoluteur lui-même.

Un tel système de traitement pourra être simplifié par l'utilisation des corrélateurs à mémoire mentionnés au paragraphe 1 ; ils permettront d'effectuer un traitement entièrement analogique dans un encombrement et avec une consommation très réduits (figure 5).

Les performances demandées au corrélateur à mémoire sont les mêmes que celles du convoluteur piézoélectrique. La durée de mémorisation reste essentiellement variable en fonction des utilisations.

Identification : Le système de traitement est analogue au précédent. La différence réside dans le fait qu'un signal issu du récepteur est mémorisé pour être comparé successivement à des répliques tenues en bibliothèque (figure 6).

Les circuits complémentaires de traitement et de décision sont identiques.

La période de renouvellement du signal à l'entrée et le nombre de répliques à lui comparer déterminent par rapport à la vitesse de traitement du convoluteur le domaine des signaux exploitables.

De même que précédemment l'utilisation d'un corrélateur à mémoire devrait permettre de déboucher sur une simplification du système.

Ce type de traitement est applicable par exemple à l'identification de la signature de certaines cibles radars.

3.3. Traitement associé à un radar à ouverture synthétique

Le principe de l'antenne synthétique appliqué à un radar à vision latérale, aéroporté ou à bord de satellite, permet d'obtenir des photographies du sol d'une haute résolution de façon pratiquement indépendante des conditions météorologiques et de la luminosité. Un tel système trouve des applications aussi bien dans le domaine civil : cartographie, étude des océans, de la végétation, détection de ressources terrestres que dans le domaine militaire : reconnaissance.

La zône de terrain visualisée par un radar à vision latérale est parallèle à la route suivie par le porteur, sa largeur peut varier entre quelques kilomètres et quelques dizaines de kilomètres. La résolution dans la direction perpendiculaire à cette route, dite "résolution en distance radar" est généralement obtenue par l'utilisation de la technique de compression d'impulsion. Cette technique permet tout en émettant un signal de puissance crête relativement faible d'avoir l'équivalent d'une impulsion très fine avec un bilan énergétique suffisant.

Le principe de l'antenne synthétique est utilisé pour améliorer de façon importante la résolution dans le sens de l'avancement du porteur dite "résolution en route" par rapport à la résolution naturelle de l'antenne du radar. Cette dernière est fonction de la distance D entre l'antenne et le point observé, elle vaut $\frac{D\lambda}{L}$ où λ est la longueur d'onde et L la longueur de l'antenne, alors que la résolution ultime qui puisse être obtenue par l'utilisation de l'antenne synthétique est $\frac{L}{2}$, indépendante de la distance D.

L'antenne synthétique est obtenue en déplaçant une source fonctionnant en émission-réception et en corrélant les signaux reçus avec une réplique.

Cette réplique est fonction du déplacement de la source. La fonction de corrélation ainsi obtenue est telle que les signaux en provenance d'un même point de l'espace se trouvent additionnés en phase. Dans le domaine spectral, on réalise ainsi le filtre spatial adapté au point considéré. L'application aux radars à vision latérale utilise un déplacement linéaire de la source constituée par l'antenne du radar.

L'antenne synthétique résultant du déplacement linéaire d'une source d'une longueur d et du traitement adéquat du signal reçu possède un diagramme équivalent à celui d'une antenne réelle de longueur 2 d.

Le traitement du signal issu d'un écho ponctuel situé à une distance D de l'antenne a une durée égale au temps d'éclairement de cet écho par le radar, qui correspond à un déplacement de l'antenne de $2D \frac{\lambda}{L}$. On en déduit la largeur de cet écho, vu par l'antenne synthétique équivalente, qui est de $\frac{L}{2}$.

Le traitement qui réalise la focalisation de l'antenne nécessite la connaissance du déphasage du signal reçu d'un écho ponctuel pendant la durée de son éclairement. Le radar utilisé est donc un radar cohérent qui conserve, pour servir de référence lors de la démodulation des signaux reçus, la mémoire de la phase du signal émis.

A ce radar est associé un système de traitement qui effectue la corrélation du signal reçu avec un signal de référence, ou réplique, contenu dans le corrélateur ou généré par lui. Les quantités d'information à traiter très importantes, et la complexité du traitement ont conduit, dans un premier temps à réaliser des unités de traitement au sol, le signal brut en sortie du radar étant enregistré en vol ou transmis au sol par liaison hertzienne.

Toutefois, il peut être intéressant, soit pour des besoins opérationnels (acquisition rapide de la carte du terrain survolé), soit surtout pour des nécessités techniques (limitation de la quantité d'informations à transmettre en particulier pour un système embarqué sur satellite) d'effectuer le traitement à bord, en temps réel. Se pose alors le problème de la faisabilité du traitement, compte tenu des contraintes sur le poids, le volume et la consommation des circuits auxquelles sont soumis les matériels aéroportés ou spatiaux. Pour des performances données, cette faisabilité est profondément liée à la technologie utilisée.

En particulier, le convoluteur piézoélectrique est bien adapté au traitement d'antenne synthétique et offre l'avantage d'une grande rapidité de calcul associée à un volume et une consommation relativement faibles.

Les principaux traitements envisageables pour un radar à antenne latérale synthétique sont actuellement de quatre types :

- corrélation optique,

- corrélation numérique,

- corrélation acoustique,

- corrélation à CCD.

Ils comportent deux aspects : d'une part la mémorisation du signal reçu sur une certaine durée, d'autre part le calcul proprement dit de corrélation (domaine temporel) ou le filtrage adapté (domaine spectral) qui permettent l'affinage de ce signal en "route". Par ailleurs, dans l'état actuel de la technologie, le traitement d'affinage en "route" autre qu'optique nécessite un échantillonnage en distance de la vidéo radar. Le signal correspondant à la zone utile de réception est ainsi découpé en cases de distance, lesquelles sont traitées de façon séquentielle ou parallèle en fonction de la nature et de l'organisation des circuits de traitement.

- Le traitement par corrélation optique s'effectue au sol à partir d'enregistrements sur film (lequel sert de mémoire) en vol. Il n'autorise pas un traitement embarqué en temps réel. Ce système a été le premier utilisé et donne d'excellents résultats [11].

- Le traitement par corrélation numérique utilise une mémoire et des opérateurs numériques pour effectuer les calculs. Plusieurs opérateurs en parallèle sont nécessaires pour traiter en temps réel le signal reçu dans les différentes cases de distance. Le nombre de ces opérateurs croît proportionnellement au nombre de cases de distance ce qui conduit assez vite à des volumes de circuits prohibitifs.

- Le traitement par corrélation acoustique utilise une mémoire qui peut être soit une mémoire numérique, soit une mémoire analogique CCD. L'emploi de cette dernière technologie dépend essentiellement de la durée de mémorisation du signal qui est typiquement de l'ordre de 1 seconde pour les radars sur avion et de l'ordre de 100 ms pour les radars sur satellite; une durée de mémorisation de l'ordre de la seconde n'est possible qu'en conditionnant les circuits à basse température ce qui constitue une contrainte importante pour un matériel aéroporté. L'opération de corrélation proprement dite est effectuée au moyen du convoluteur piézoélectrique qui traite, case de distance après case de distance les informations stockées en mémoire. L'intérêt de ce dispositif est qu'il permet le calcul de la fonction de corrélation par morceaux et non pas seulement point par point comme le corrélateur numérique.

- Le corrélateur à CCD intègre la mémorisation et le calcul dans un même circuit. Il se compose en effet, pour chaque case de distance, d'un filtre transversal capable de traiter un signal complexe (soit quatre filtres transversaux réels) dont les coefficients, représentatifs de la fonction de référence sont inscrits dans le dispositif. En ce qui concerne les applications aéroportées, ce système comporte deux inconvénients majeurs :

- durée importante de la mémorisation qui nécessite un conditionnement à basse température

- diversité des fonctions de référence qui dépendent de la distance radar; ceci nécessite plusieurs types de filtres associés chacun à une ou quelques cases de distance.

La comparaison qui suit est faite entre les deux systèmes de traitement qui offrent actuellement la plus grande souplesse d'emploi à savoir : le traitement numérique et le traitement acoustique fondé sur l'utilisation du convoluteur piézoélectrique.

Corrélateur numérique

Un schéma synoptique du système de traitement par corrélation numérique est fourni figure 7.

Le signal vidéo issu du récepteur après démodulation cohérente est échantillonné et converti sous forme numérique dans la zone utile de chaque période de récurrence. Cet échantillonnage découpe la bande de terrain visualisée en un nombre de cases de distance, variable suivant la largeur de la bande et la résolution recherchée en distance radar. Le signal reçu dans chaque case de distance est ensuite mémorisé après un éventuel pré-traitement.

Celui-ci permet, quand c'est possible, de diminuer le nombre d'échantillons à traiter après avoir comprimé la bande de fréquence du signal. Les échantillons mémorisés représentent soit l'amplitude et la phase (ρs, θs), soit les composantes cartésiennes (xs, ys) du vecteur signal.

La dimension de la mémoire dans le sens de la route (nombre m de récurrences mémorisées) dépend de la largeur du signal de référence utilisé dans l'opération de corrélation, elle-même définissant la résolution "en route."

La résolution optimale L/2 suppose qu'est traitée toute l'information contenue dans le lobe principal de l'antenne (limité à l'ouverture à 3 dB), ce qui correspond à

$$M = \frac{\theta o \ D}{V} \ fr$$

où θo = largeur à 3 dB du diagramme d'antenne

  D = distance radar-cible dans l'axe de l'antenne

  V = vitesse du porteur

  fr = fréquence de récurrence du radar.

La résolution r obtenue en ne considérant que m échantillons est en première approximation égale à :

$$r = \frac{L}{2} \cdot \frac{M}{m}$$

Lors de chaque période de récurrence, l'opération effectuée est une corrélation discrète entre le signal mémorisé dans chaque case de distance et un signal de référence. Ce signal de référence possède les caractéristiques d'amplitude et de phase (ρ_r, θ_r) du signal reçu d'un écho ponctuel fixe idéal, à savoir une amplitude variable en fonction du diagramme de rayonnement de l'antenne et une phase quadratique :

$$\theta_s = \phi_o + Kt^2 \ (t = o \ \text{dans l'axe de l'antenne})$$

où

$$\phi_o = \frac{4\pi D}{\lambda}$$

$$K = \frac{2\pi \ V^2}{\lambda \ D}$$

soit sous forme échantillonnée $\theta_s (i) = \phi_o + KT^2 i^2$

où T est la période de récurrence.

L'opération de corrélation s'écrit :

$$S (n) = \Sigma_i \ \rho_s (i-n) \cdot \rho_r (i) \ \exp \left\{ j \left[\theta_s (i-n) - \theta_r (i) \right] \right\}$$

la sommation s'étendant sur les m valeurs mémorisées.

Le calcul de S(n) est effectué pour toutes les cases de distance au cours de chaque période de récurrence. Le signal affiné S(n) est ainsi élaboré de façon discrète pour chaque case de distance. Seule l'amplitude de S(n) est prise en considération.

Le calcul de chaque point de la fonction de corrélation avec des signaux complexes nécessite théoriquement de l'ordre de 4 m multiplications et 4 m additions, (m pouvant atteindre quelques centaines).

L'utilisation de quantifications particulières de l'amplitude et de la phase du signal permet de s'affranchir des multiplications qui sont longues à effectuer.

La durée de calcul d'un point de la fonction de corrélation pour une case de distance devient alors typiquement 50 ns x m en utilisant de la logique rapide (ECL).

L'opérateur numérique câblé capable de ce calcul et utilisant une quantification de phase en 4 éléments binaires occupe un volume de circuits d'environ 1 litre pour une consommation de l'ordre de 15 W.

L'utilisation d'une quantification de phase plus fine fait croître le volume et la consommation (en première approximation doublement à chaque fois qu'un élément binaire est ajouté).

Compte tenu du nombre de cases de distance à traiter et de la fréquence de récurrence du radar, il est possible d'en déduire le nombre d'opérateurs nécessaires pour effectuer le traitement en temps réel et donc un volume et une consommation approximative de la partie calcul.

Les lois de référence dépendent théoriquement de la distance D; toutefois il est possible avec une bonne approximation d'utiliser la même loi pour un ensemble de cases de distance. Ces lois sont, soit stockées en mémoire, soit générées dans le calculateur; cette dernière solution est plus souple d'emploi et offre la possibilité d'introduire, au niveau de la référence des termes correctifs pour tenir compte de désadaptations éventuelles introduites en amont dans le traitement.

Corrélateur acoustique

Un schéma synoptique du système de traitement par corrélation acoustique est fourni figure 8.

Le corrélateur acoustique est fondé sur l'utilisation du convoluteur piézoélectrique associé à une mémoire numérique.

La première partie du traitement jusqu'à la mise en mémoire, est similaire à celle du traitement numérique, à savoir échantillonnage, conversion analogique-numérique, prétraitement, mémoire.

Ceci n'est qu'un exemple d'utilisation, l'emploi du convoluteur n'implique pas nécessairement que toute cette partie soit traitée en numérique. Les prétraitements et la mémorisation peuvent être réalisés au moyen de dispositifs CCD avec les réserves déjà mentionnées quant au temps de mémorisation. Les interfaces mémoire-convoluteur sont alors différentes de celles associées à une mémorisation numérique.

La deuxième partie du traitement consiste à effectuer l'opération de corrélation entre le signal reçu mémorisé (sous la forme amplitude-phase ρ_s, θ_s ou sous forme cartésienne x_s, y_s) et le signal de référence.

Le convoluteur reçoit à l'une de ses entrées le signal lu dans une case de distance de la mémoire après mise en forme. L'autre entrée reçoit, en synchronisme, le signal de référence obtenu de la même manière (lecture de mémoire et interface) ou généré à partir d'un oscillateur local balayé en fréquence. Le produit de corrélation est obtenu en sortie après détection et échantillonnage.

La durée maximale T du signal acceptable par le convoluteur est de 12 µs.

La mémoire relative à une case de distance est remplie au fur et à mesure de l'avancement de l'avion. L'information mémorisée a une durée correspondant au temps d'éclairage d'une cible ponctuelle soit environ 1 seconde. On réalise donc une compression de temps de l'ordre de 10^5 tout en conservant le produit BT qui vaut $\frac{2 \theta_0^2 D}{\lambda}$. Cette compression de temps nécessite une mémoire tampon à lecture très rapide puisqu'il faut relire en 12 µs un nombre m d'échantillons de l'ordre de quelques centaines.

Une solution envisagée pour générer le signal de référence est d'utiliser un oscillateur local balayé linéairement en fréquence, centré à la fréquence de travail du convoluteur. Une compression correcte du signal à une distance donnée, impose à la référence d'être à la bonne fréquence centrale, d'avoir une caractéristique de pente déterminée et parfaitement linéaire.

Des écarts par rapport aux spécifications théoriques se traduisent par un décalage en temps de la fonction de corrélation en sortie, équivalent à un décalage "en route", à une perte sur l'amplitude et à un élargissement du signal affiné.

Il est donc nécessaire d'une part, d'asservir la fréquence centrale de l'oscillateur local à la fréquence de transposition du signal issu de la mémoire, d'autre part d'asservir sa pente.

Les précisions nécessaires sont de l'ordre de 10 KHz pour la fréquence centrale et de la dizaine de degrés sur les caractéristiques de phase.

L'intérêt du convoluteur piézoélectrique réside en ce qu'il fournit non pas un point, mais la fonction de corrélation. Sous réserve de mémoriser le signal reçu sur une durée supérieure à celle du signal de référence cette technique conduit à calculer la fonction de corrélation par morceaux. Ceci évite d'avoir à effectuer le calcul lors de chaque période de récurrence pour toutes les cases de distance. Il est ainsi possible avec un seul convoluteur de traiter un grand nombre de cases de distance et de couvrir un domaine plus important, à volume de

traitement identique, qu'avec des opérateurs purement numériques.

Le signal corrélé, après détection est échantillonné et mémorisé de façon à regrouper, pour une case de distance les différents morceaux calculés séparément. Ce signal de sortie peut être alors exploité. Il représente directement la carte radar affinée du sol.

Le convoluteur acoustique avec ses interfaces associées : modulateurs, amplificateurs, filtres, occupe un volume d'environ 1 litre pour une consommation de l'ordre de 5 watts.

La génération du signal de référence avec ses boucles d'asservissement a un volume et une consommation analogues.

Comparaison numérique-acoustique

Le tableau 5 présente quelques caractéristiques d'un système de traitement par corrélation associé à un radar à antenne latérale synthétique aéroporté dans deux configurations :

- visualisation d'une bande de 5 km de large avec une résolution de l'ordre de 5 m x 5 m
- visualisation d'une bande de 50 km de large avec une résolution de l'ordre de 10 m x 10 m.

	Zone de largeur 5 km	Zone de largeur 50 km
Distances limites	5 km - 10 km	50 km - 100 km
Résolution	5 m x 5 m	10 m x 10 m
Nombre de cases de distances	2 000	10 000
Nombre d'échantillons mémorisés en route (valeur moyenne)	40	90
Nombre d'opérateurs numériques	2	8
Nombre d'opérateurs acoustiques	1	1
Taille mémoire du corrélateur numérique	80 K mots	900 K mots
Taille mémoire du corrélateur acoustique	120 K mots	1 400 K mots

Tableau 5

L'examen de ces caractéristiques montre que du point de vue volume et consommation l'utilisation du traitement par corrélation acoustique est d'autant plus intéressante que le nombre de cases de distance à traiter s'accroît. Il faut noter que les tailles de mémoire sont peu différentes dans les deux systèmes avec une mémoire un peu plus importante pour le corrélateur acoustique, à cause de l'accroissement du nombre de récurrences mémorisées pour le calcul par morceaux de la fonction de corrélation et du stockage du signal reçu pendant la durée de ce calcul.

L'utilisation du convoluteur piézoélectrique offre par rapport aux techniques purement numériques d'autres avantages. Il permet de s'affranchir des effets d'une quantification grossière sur la phase des signaux, quantification utilisée dans les opérateurs numériques décrits et qui dégrade la qualité de la fonction de corrélation : perte d'amplitude du pic de corrélation, remontée des lobes secondaires et de l'énergie contenue dans le piédestal. Les performances obtenues avec le convoluteur piézoélectrique sont cependant limitées par les quantifications nécessaires pour la mise en mémoire des informations.

La dynamique du convoluteur (40 à 50 dB) est nettement suffisante pour une fonction de cartographie. Les opérateurs numériques qui ont servi de comparaison possèdent une dynamique de 40 dB; l'avantage des circuits numériques est qu'il est aisé, en augmentant le nombre d'éléments binaires, d'accroître la dynamique, au détriment cependant de la complexité et du volume.

Les différents signaux parasites inhérents au fonctionnement du convoluteur : autoconvolution ou convolution trajet double, ne perturbent pas le fonctionnement du dispositif puisque le signal de sortie est échantillonné au moment de l'apparition du pic de corrélation, et qu'un intervalle de temps de l'ordre de T (12 µs) est prévu entre le traitement de deux cases de distance successives.

Une des difficultés d'emploi du convoluteur est qu'il suppose un échantillonnage et un codage rapide (quelques dizaines de nanosecondes) du signal de sortie après détection. Une

compression logarithmique de l'amplitude de ce signal diminue le nombre d'éléments binaires de quantification (on se limite à 5 éléments binaires soit 30 dB de dynamique) et facilite la réalisation du convertisseur analogique-numérique.

Enfin, le convoluteur piézoélectrique dans cette application travaille en mode synchrone : il *faut entrer simultanément le signal reçu et le signal de référence.* L'utilisation d'un corrélateur à mémoire qui conserve la mémoire du signal de référence, supposé être le même pour un ensemble de cases de distance, simplifierait sensiblement le matériel.

CONCLUSION

Le convoluteur piézoélectrique est l'un des composants à ondes acoustiques de surface qui par ses performances peut remplacer avantageusement certains systèmes numériques utilisés jusqu'ici dans les traitements de corrélation.

Il est particulièrement adapté au traitement rapide des signaux émis par les radars modernes, à produit B.T important, de plusieurs centaines, et à bande passante large, de plusieurs dizaines de mégahertz.

Son utilisation est rendue compatible de traitements numériques ou analogiques complémentaires par des interfaces mémoires et convertisseurs, indispensables pour adapter les échelles de temps des signaux et leurs formes.

Pour ces traitements l'ensemble convoluteur-interfaces offre alors par rapport aux techniques purement numériques des perspectives d'amélioration sur les vitesses de calcul, l'encombrement, la consommation et le coût.

Dans ce domaine de la corrélation appliquée aux signaux radars à large bande, l'intérêt se porte maintenant vers les corrélateurs à mémoire qui pourront remplacer le convoluteur piézoélectrique dans certains systèmes de traitement.

Leur utilisation , encore du domaine du laboratoire, devrait conduire à des systèmes extrêmement simplifiés dans leur mise en oeuvre, en particulier au niveau des interfaces et de l'exploitation qui restent un point majeur lié aux dispositifs acoustiques.

REFERENCES

1 – Paiges, E.G.S., "Dispersive filters : their design and application to pulse compression and temporal transformation", Proc. Int. Seminar on Component Performance and System Applications of SAW Devices (Avemore, Scotland), pp 167 – 180, 1973, IEE Publication 109.

2 – G.R Nudd. "Transform domain signal processing using surface wave chirp filters". Compte rendu réunion internationale DRME sur les applications des composants micro-ondes acoustiques (Paris 1976) pp 7-22.

3 – H. Gautier, C. Lardat, C. Maerfeld, "Convolution et corrélation à l'aide de composants acoustiques à ondes de surface", Colloque GRETSI (Nice 1977) pp 57/1 – 57/6.

4 – P. Defranould, H. Gautier, C. Maerfeld, P. Tournois, "P-N diode memory correlator", Ultrasonics Symp. Proc. pp 336-347, 1976.

   H. Gautier, C. Maerfeld, P. Tournois, "Convolution et Corrélation à mémoire au moyen de dispositifs acoustiques à ondes de surface", AGARD Conf. Proc., on Impact of CCD and SAW Devices on Signal Processing and Imagery in Advanced Systems, 3.7., 1977.

5 – Ingebrigtsen, K.A., et al, "A Schottky Diode Acoustic Memory and Correlator", Appl. Phys. Lett., vol. 27, p.170, 1975

6 – E. Stern, "Analog memory correlators for radar signal processing", AGARD Conf. Proc. 197, on New Devices, Techniques and Systems in Radars pp 9/1 – 9/14, 1976.

7 – P. Defranould, C. Maerfeld, "Acoustic Convolver Using Multistrip Beam Compressor" 1974, Ultrasonic symp. Proc. p224, IEE New york, 1974.

8 – D.P. Morgan, J.M. Hannah, "Programmable Correlation Using Parametric Interactions in Acoustic Surface Waves", Electron". Lett., Vol 8, p 40, 1972.

9 – Smith, J.M. et al, "Surface Acoustoelectric Convolvers", 1973 Ultrasonic Symposium Proceedings, (IEEE New-york. 1973) p 142.

10 – Kino, G.S., et al, "Signal Processing by Parametric Interactions in Delay Line Devices", IEEE Trans. Sonics and Ultrasonics, Vol SU20, p.162, 1973

   P.G. Borden, G.S. Kino, "Correlation with the storage convolver", Appl. Phys. Letts., 29-9, pp 527-529, 1976.

11 – J. Genuist. "Radar à faisceau latéral utilisant une antenne synthétique", AGARD Conf. Proc. 197 on New Devices, Techniques and Systems in Radar pp. 17/1 – 17/15, 1976.

DISCUSSION

Roberts

In the synthetic aperture processor you have a video store which is digital. How many bits do you think the words have to be and how many different reference chirps do you require for different ranges?

Réponse d'Auteur

Le signal vidéo est converti sous forme numérique en 7 bits pour chacune des composantes. La mise en mémoire s'effectue après un codage particulier sur un nombre inférieur de bits. Pour le traitement numérique associé au premier exemple (zone de 5 km) il faut environ 30 lois de références différentes.
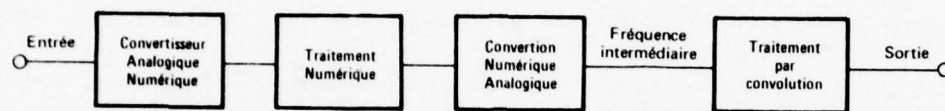
Figure 1   BLOC DIAGRAMME DE TRAITEMENT
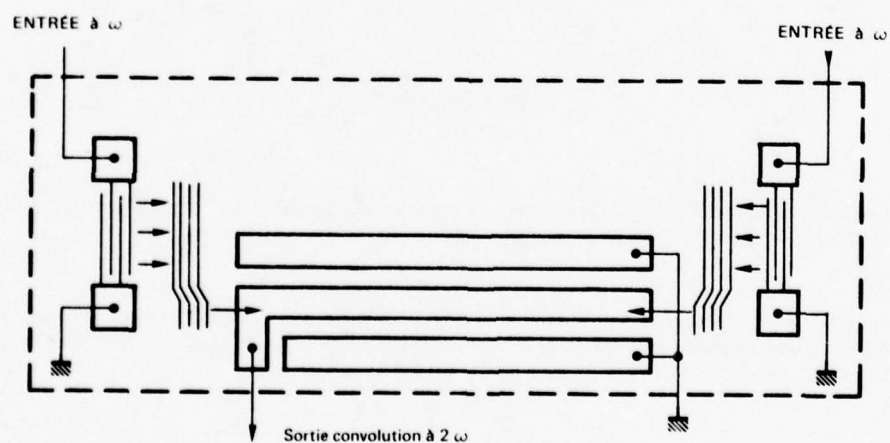
Figure 2   SCHÉMA DU CONVOLUTEUR PIÉZOÉLECTRIQUE
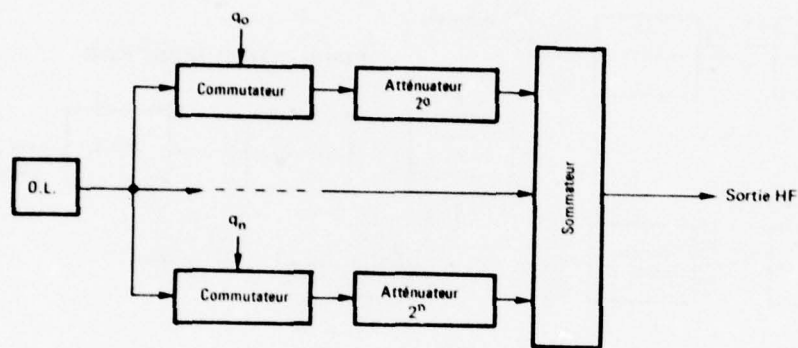
MODULATION D'AMPLITUDE
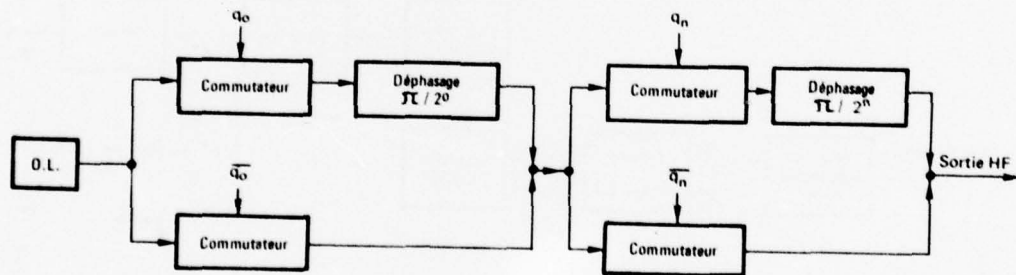
MODULATION DE PHASE

Figure 3   MODULATEURS
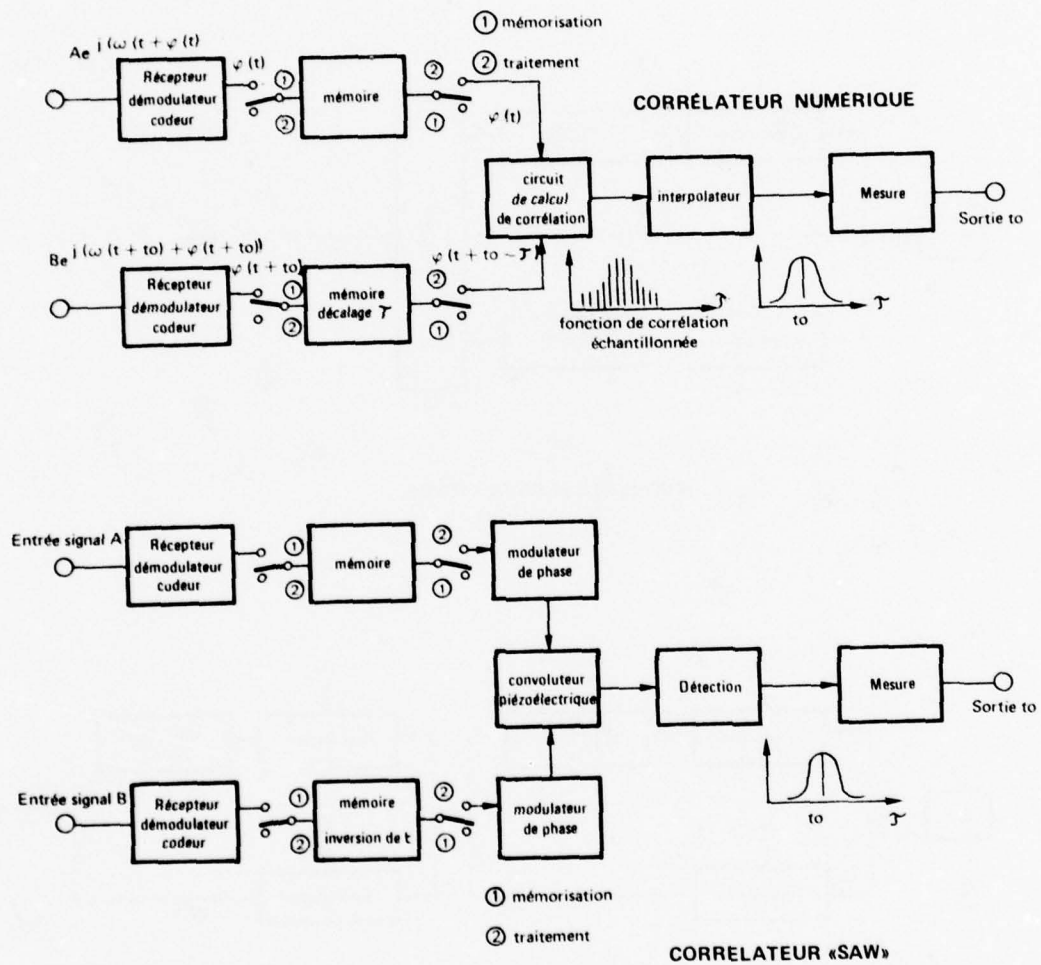
Figure 4   MESURE DE TEMPS PAR CORRELATION

Figure 5   TRI DE SIGNAUX PAR CORRÉLATION

ENTRÉE

| Réception | synchronisation | | Bibliothèque des répliques |

| Normalisation des signaux |

| échantillonnage quantification | mémoire | interface | | interface |

convoluteur

traitement décision

SORTIE

VERSION ACTUELLE : TRAITEMENT NUMÉRIQUE

CORRÉLATEUR PIÉZOÉLECTRIQUE

ENTRÉE

| Réception | synchronisation | circuit d'écriture | | Bibliothèque des répliques |

| Normalisation des signaux | Signal à mémoriser | corrélateur à mémoire | interface |

traitement décision

SORTIE

VERSION FUTURE : CORRÉLATEUR A MÉMOIRE

Figure 6   IDENTIFICATION DE SIGNAUX PAR CORRÉLATION
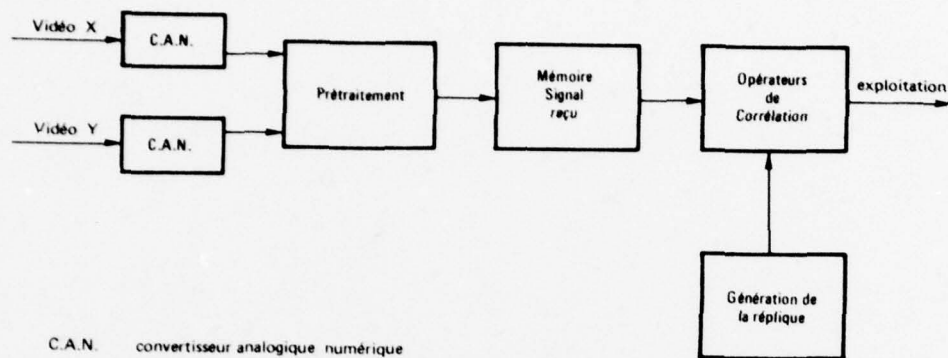
Figure 7   CORRÉLATEUR NUMÉRIQUE

C.A.N.   convertisseur analogique numérique
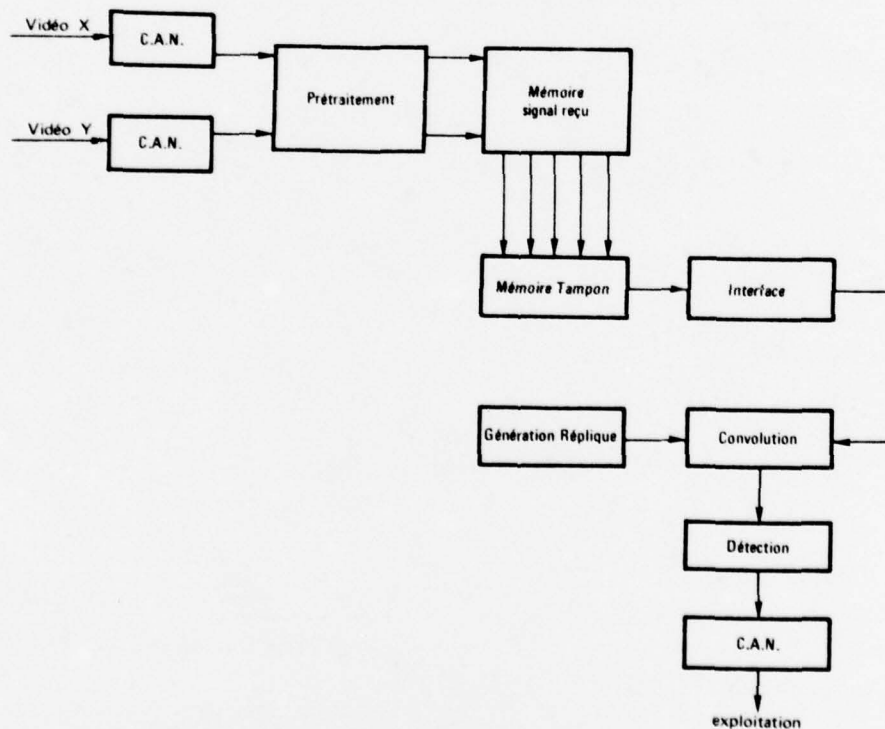
Figure 8   CORRÉLATEUR ACOUSTIQUE

OPERATION OF SAW REFLECTIVE ARRAY PULSE COMPRESSORS IN A HIGH PERFORMANCE RADAR
WITH ~0.4 METER RANGE RESOLUTION[*]

W. J. Ince
Massachusetts Institute of Technology
LINCOLN LABORATORY
Lexington, Massachusetts 02173

S. C. Cottrill
P. Gelznis
RCA

Kiernan Reentry Measurements Site
Kwajalein, Marshall Islands

SUMMARY

The application of L-band reflective array pulse compression devices with a time-bandwidth product of 5120 to a high resolution radar is described. This is believed to be the largest time-bandwidth product for any reflective array compressor (RAC) in an operational system.

The subject radar is the Lincoln Laboratory ALCOR, which is a C-band instrumentation tracking radar designed to provide precision trajectory and signature information on space vehicles. In its waveform inventory ALCOR has a wideband operating mode with a 512-MHz instantaneous bandwidth (0.4 meter range resolution) that permits resolution of the individual scattering centers of a target.

Since the original construction of ALCOR, techniques have become available for synthesizing a matched filter to the wideband chirp transmitted waveform utilizing SAW devices. Two receiver channels, matched to left- and right-circular polarization respectively, have been implemented. An earlier (successful) attempt at a matched filter design, utilizing a band-partitioned approach rather than SAW technology, required 32 parallel subchannels each handling only 16 MHz of bandwidth. In this prior work only a single matched filter channel, to process either LC or RC polarization, was implemented due to the formidable number of components required.

The SAW approach utilizing RAC devices permits a very considerable savings in equipment, and the capability of providing amplitude weighting and phase compensation within the RAC itself. RAC devices have operated reliably in the ALCOR radar for approximately four years.

INTRODUCTION

ALCOR is a high (range) resolution, pencil beam, monopulse tracking radar. Designed as a range instrumentation system, it has been operating for a number of years collecting trajectory and signature data on satellites and reentry objects.

The radar has three basic modes; a wideband (512 MHz) skin tracking mode that permits observation of scattering centers over a very limited range extent (90 meters); a narrowband (6 MHz) skin tracking mode whose primary function is to provide target acquisition for the wideband mode; a beacon tracking mode for target identification. All three modes are also used, in combination, for gathering trajectory data. The radar timing, selection of tracking waveforms, tracking algorithms, and data recording are all under the control of a digital computer. The ALCOR transmission in either of the two skin tracking modes is a linear FM pulse of approximately 10 microseconds duration. The radiation is right circularly polarized; on receive, both left and right circular polarizations are processed.

During the initial design of ALCOR in 1964 it was clear that the amount of WB data recorded would have to be severely limited in order to stay within the capabilities of digital magnetic recording equipment that was commercially available at that time. Furthermore, the capability to construct a matched receiver from components having the full instantaneous bandwidth of the WB mode did not exist. The decision was taken to base the design of the receiver signal processing for the WB mode on a time-bandwidth exchange scheme ("stretch") whereby the instantaneous signal bandwidth is reduced to that of the NB mode; 6 MHz. Correspondingly, in the stretch processor the time axis is expanded by 85:1. Hence it was possible to multiplex the WB and NB signals through the same receiver channel and to use the same pulse compression network, weighting filter and transversal equalizer. The major components of the ALCOR receiver are illustrated in Figure 1.

The recorded radar data comprises digital video samples of logarithmic amplitude and phase, for both LC and RC channels. The recording windows for the NB and WB modes, assuming a 10 MHz sampling rate, are 2.5 km and 30 meters respectively. In the WB mode, a 30 meter window size was originally considered adequate to fully cover the extent of any space object of interest. Subsequently a capability of recording over three 30-meter contiguous range windows, to provide a 90-meter recording window, was added. This capability was obtained at the cost of a considerable increase in the receiver hardware, in the form of two additional correlation mixers in the stretch processor and a number of delay lines in the multiplexer. Currently, the total data rate for the 90-meter window in the WB mode exceeds the recording capacity of the system.

Table 1
ALCOR SYSTEM PARAMETERS

Frequency	~5.4 – 5.9 GHz
Antenna beamwidth	0.3 deg.
Peak power	4 MW
Pulse length	10 microsecs
Modulation (skin modes)	Linear FM
Modulation bandwidth	512 MHz (WB mode)
	6 MHz (NB mode)
PRF	20–350 Hz
Sensitivity (skin modes)	23 dB S/N ratio on a $1m^2$ target at 1000 km range
Recording capability	280 Kbytes per sec (8 bit bytes)
Video Sampling (10 MHz Rate)	NB; 170 samples amp. and phase over a 2.5 km range window
	WB; 170 samples amp. and phase over each of three 30m range windows
Quantization	Amplitude 7 bits; Phase 8 bits

From this discussion it will be evident that the original ALCOR WB signal processing and recording architecture are not easily adaptable to the gathering of data on extended target complexes, and an alternative approach was adopted. Additional parallel LC and RC receiver channels with SAW analog matched filters were implemented, enabling targets anywhere within the PRI to be processed. Secondly, in order to overcome the recording bottleneck it was mandatory to undertake post processing data condensation prior to recording. This sorting and filtering of the data is carried out, after conversion to 250 MHz baseband, in high speed special purpose digital logic circuitry.

THE ALL-RANGE WIDEBAND CHANNEL

The new ALCOR processing channels, known collectively as the All Range Wideband Channel (ARWBC) are illustrated in Figure 1. In order to avoid overburdening the ALCOR system, additional independent recording and displays were implemented to handle the ARWBC data. Due to the data condensation, only modest recording capacity is required.

The ARWBC target information recorded consists of

a) Target reports on up to 30 targets each PRI

b) Twenty consecutive digital samples at a 400 MHz rate on a single selectable target per PRI

A supplementary, but nonetheless important, role of ARWBC is to provide an extended range acquisition window for the ALCOR WB tracker. Normally in ALCOR, target designation and lock-on requires manual intervention. Since the WB tracker is very responsive, this necessitates considerable dexterity on the part of the operator. The wide acquisition window provided by the ARWBC display materially aids this process. All targets exceeding the threshold show up on the display although data is recorded on only 30 detected targets. In addition, target ranges can be supplied to the ALCOR computer for automatic acquisition.

An early version of the ARWBC matched filter, the so-called All Range Processor (ARP), was implemented in 1972. It has since been superceded by a RAC design. Figure 2 illustrates the tremendous reduction in physical hardware achieved with the RAC configuration. The figure shows a dual polarization matched filter assembly of two RAC's housed individually in thermostatted ovens, and the four cabinets of solid state components that would be required to achieve the equivalent electrical performance using a channelized receiver approach. In order to convey an appreciation for the simplicity of the RAC configuration versus the ARP, a brief description of the latter will be given.

THE ALL-RANGE PROCESSOR

The principle of operation of the ARP is illustrated in Figure 3. The received signal is partitioned in the frequency domain into 32 subchannels, each having a bandwidth of 16 MHz and a time-bandwidth product of 5. Pulse compression is obtained with the combination of a dispersive bridged tee network and a fixed delay in each subchannel.

The ARP can process only one polarization at a time. A polarization switch at the RF inputs enables either LC or RC polarization to be selected. The ARP accepts the C-band ALCOR signal and, after transversal equalization, divides it into four paths. The four components are down converted to a common center frequency of 768 MHz. Each of the four paths are further divided by eight. In each of the 32 paths so formed the signal is mixed down to an IF center frequency of 40 MHz where pulse compression is performed. The compressed and delayed outputs are mixed back up to 768 MHz and combined into a single output. After combining, the signal frequency is translated up to C-band where amplitude weighting is performed in an interdigital filter. Finally, the compressed, weighted ARP output signal is mixed down to the log IF frequency of 1 GHz.

When the RAC system became available, the ARP was maintained on line as a spare receiver channel in case of catastrophic RAC failure, an event which never actually occurred. With the reliability of the RAC devices proven, the ARP was eventually deactivated.

The total complement of ARP hardware, including mixers, frequency sources and power supplies, which occupies seven equipment racks, is shown in Figure 4. A pair of RAC's replaced five of the seven equipment cabinets. The transversal equalizers were retained for operation with the RAC devices. L-band mixers and log amplifiers were also required.

REFLECTIVE ARRAY COMPRESSOR DESIGN

The design, fabrication and properties of reflective array compressors of the type used in the present work have been described extensively by Williamson et al[2-4]. Devices were fabricated on LiNbO$_3$ substrate material with the topography illustrated in Figure 5.

The RAC design center frequency of 1 GHz was selected after taking into consideration three competing factors; insertion loss and fabrication difficulty, both of which are inversely proportional to frequency; and desire for a large fractional bandwidth, which suggests a high frequency. The special problems[5] encountered at frequencies as high as 1 GHz, associated with dimensional tolerances, attenuation and dispersion of the medium, have been successfully overcome in the devices supplied.

The surface wave reflection grating consists of a "herringbone" array of ion-beam etched grooves of width 1.2 μm and variable depth. The groove spacing increases linearly with distance from the input-output transducer pair. The component of the chirped signal corresponding to a given frequency is reflected from a small region of the grating where the groove spacing matches the wavelength of the propagating surface wave. The geometry is such that a wave launched in the Z direction is reflected through 90° by the grating in the path of the wave. A second grating, which is a mirror image of the first grating about the Z symmetry axis, introduces another 90° reflection to reverse the direction of propagation of the surface wave towards the output transducer, where it is intercepted. The acoustic beam is approximately 100 wavelengths wide.

The reflection grating pattern for each RAC was reproduced from a master photomask. The grating pattern was first copied onto the substrate in photo-resist and then ion-beam etched to form the grooves of the reflection grating. The grating reflectivity is a function of the groove depth, which was varied progressively along the length of the grating, ranging from 20 to 400 Å, to compensate for the variation in RAC characteristics with frequency and to achieve an approximately Hamming amplitude-weighted response in the finished product.

The input and output transducers consist of 5 fingers, each 0.8 micron wide. The transducer pattern was produced by electron-beam lithography and replicated by photolithographic techniques.

A packaged RAC device is shown in Figure 6. Portions of the reflection grating are visible on the substrate. In order to decrease input-output direct feedthrough, a metal septum has been placed between the transducers, which are separated by only 3 mm. The combination of the transducer impedance and the surface wave impedance are highly reactive, and a broadband input match is difficult to achieve. Therefore, only a modest impedance transformation was attempted, to reduce the mismatch loss while retaining a relatively frequency insensitive power transfer characteristic. Connection to each transducer is made from a 50 ohm microstrip transmission line with an inductive fine gold wire which partially compensates the transducer capacitance. Attenuator pads were placed at the input and output terminals of each RAC to keep the VSWR below 2:1.

Since the RAC electrical characteristics are temperature sensitive, each device has been designed to match the ALCOR transmitted waveform at an elevated temperature (~50°C), allowing the RAC's to be housed in individual ovens for temperature stabilization.

RAC OPERATING CHARACTERISTICS

A total of nine RAC devices, designed and fabricated by the Surface Wave Technology Group of Lincoln Laboratory, have been delivered to the Kwajalein field site and operated over a period of four years. The device characteristics are listed in Table II. Experience has been gained with all units. The initial installation of units 340 and 383 in the ALCOR All-Range Wideband Channel took place in October 1973.

The RAC operating temperature must be stabilized to within relatively tight limits in order to minimize variations in electrical characteristics. The most sensitive parameter is the group delay at the center frequency, the temperature dependence of this parameter being given by

$$\frac{1}{\tau_o} \frac{d\tau_o}{dTemp} = \gamma_o \left[1 + \frac{f_o T}{B\tau_o} \right] ,$$

where γ_o is the temperature coefficient of delay (+9.0 x 10^{-5} per deg C), f_o is the center frequency (1 GHz), T is the pulse width (10 microsecs), B is the instantaneous bandwidth (512 MHz) and the nominal group delay, τ_o, is 7.83 microseconds. The evaluated change in group delay is 1.1 nanoseconds per degree C. Hence, a temperature stabilization range of about ±0.1°C will maintain the delay to a small fraction of the compressed pulse width.

The RAC devices were designed to be matched to the ALCOR WB frequency ramp at an operating temperature of 50°C. Small deviations in material characteristics and surface conditions from crystal to crystal cause the oven temperature setting required for the correct chirp slope to be different from the nominal value, but typically within 10°C. The fractional change of chirp slope per degree centigrade is equal to $-2\gamma_o$ (i.e., -1.8×10^{-4} per deg C). The most significant effect of a small slope mismatch is a degradation in the time sidelobes. If Φ_q is the quadratic phase error at the frequency band edge, then

$$\frac{d\Phi_q}{dTemp} = \frac{\pi}{2}\gamma_o TB \quad.$$

$$= 41.5° \text{ phase per degree Centigrade.}$$

From the work of Klauder et al[6] a quadratic phase error of 41.5° would have only a very small effect on the compression gain and the sidelobe level. Thus the phase error produced by a ±0.1°C temperature variation would have a completely negligible effect. It is noteworthy that the thermostatted RAC ovens are capable of maintaining the temperature setting to within ±0.05°C.

Table II shows the calculated characteristics of the compressed waveform for an ideal linear FM input. The calculated sidelobe level is based upon the measured device CW insertion loss and phase versus frequency. For the first four units the sidelobe level is considerably higher than one might expect for a Hamming weighted structure. Both amplitude errors in the weighting and phase errors contribute to this problem. It was discovered that the grating grooves were unintentionally deepened by an acid solution used during a cleaning process following ion-beam etching. This had the effect of modifying the grating reflectivity such that the amplitude response versus frequency becomes skewed, as illustrated by several of the amplitude characteristics shown in Figure 7a. Both units 489 and 490 exhibit a skewed amplitude response, yet exhibit much lower sidelobes than, for example, unit 383. Since the deviations from quadrature phase response for units 489 and 490 are very small (1.5° rms), careful control of the phase response appears to be the dominant reason for the superior performance of these two devices. Figure 7b shows the compensated phase responses for the same units as for Figure 7a.

The typical phase deviation from the ideal quadratic response is generally excessive for the ALCOR application, and phase compensation has to be applied, and phase compensation has to be applied individually to each device. The typical uncompensated and compensated phase responses for a RAC are shown in Figure 8. Williamson[5] has attributed the slowly varying phase component, which is the major problem, as being due to "free surface" dispersion. Apparently the elastic properties of the surface layer of the LiNbO₃ substrate, to a depth of about 30Å are slightly different from those of the deeper material. Thus the local surface condition causes a small spread in the velocity of the surface wave. The additional small periodic phase error observable in Figure 8 is believed to be due to a periodic error in the location of the reflection grating grooves.

An internal phase compensation scheme, based upon the retardation of the surface wave by a metal film, has been used to reduce the rms phase deviation from quadratic to a few degrees. The metal film is deposited on the substrate between the two gratings, as shown in Figure 5. The width of the conducting film as a function of distance Z is determined separately for each RAC, knowing the required phase correction as a function of frequency and assuming an empirically determined relationship for the delay per unit length of conducting film. It is significant that the residual cubic and higher order phase errors contribute primarily to the near-in sidelobes which can be reduced by transversal equalizers.

Measurements of time sidelobe levels with the RAC device connected into the radar system show a general degradation from the calculated performance, due to the non-ideal characteristics of the ALCOR WB ramp generator and to internal reflections in the system. Sixteen-tap transversal equalizers are provided in the LC and RC channels to suppress sidelobes occurring within ±16 nanoseconds of the main response. The transversal equalizers should enable the residual sidelobe level within this range to be suppressed to below the -30 dB level.

Experimental estimates of the maximum sidelobe level may be a little low due to the nature of the measurement set-up although there is a high confidence that the peak sidelobe level did not exceed -30 dB in almost all cases. Measurements were made from a video display at the output of the log amplifier which follows the RAC. There is evidence that the amplifier does not respond logarithmically to the sidelobe pulse, which is only 2 nanoseconds wide. The mainlobe of the response, however, is closer to 4 nanoseconds wide between nulls, and reproduction of the mainlobe pulse is more faithful.

Figure 9 illustrates the performance of Unit 383 in the ARWBC. Figure 9a shows a sampling oscilloscope trace of the compressed carrier at the output of the RAC. Figure 9b shows the detected output of the log IF amplifier. The maximum sidelobe level is -30 dB. The X-axis scale is 5 nanoseconds per cm.

An interesting characteristic property of the RAC device is a very low sidelobe level in the region well removed from the peak of the compressed pulse. This is an additional sidelobe suppression inherent in the RAC, over and above that due to the Hamming weighting designed into the structure. It is a function of the RAC geometry and is related to the fact that the surface wave undergoes two cascaded reflections within the device.

Units 475 and 477 differ significantly in characteristics from the other devices, believed to be due to exposure of the LiNbO₃ substrate to excessive temperatures during one of the fabrication stages. This caused local variations in the substrate surface characteristics, causing the surface wave propagation to become non-uniformly dispersive. One consequence is an abnormally high deviation from the ideal quadratic phase response, which was difficult to compensate. With device No. 475, for example, the residual rms deviation is 7.5°. In addition, the chirp slope differed substantially from the design value, requiring a low operating temperature to match the ALCOR WB ramp. Since the required temperature was out of the range of the oven controller the chirp bandwidth was, instead, reduced by 900 kHz. The impact on the ALCOR system performance was negligible. When matched simultaneously to the new ramp slope the revised operating temperatures of SN475 and 477 were 40.1°C and 45.7°C respectively.

TABLE II. Characteristics of RAC Devices Used in ALCOR All-Range Wideband Channel

Device Number	Calculated 3dB Pulse width*(nsec)	Calculated Peak Sidelobe Level (dB)for Ideal Ramp. Near** (≤10 nsec)	Far (>10nsec)	Measured Highest Sidelobe (dB). RAC Installed in ARWBC No TE's	With TE's	RMS Deviation from Quadratic Phase (Deg.)	CW Insertion Loss (dB)	Direct Feed-through Level (dB)	Operating Temp.†† (Deg.C)
340	2.4	-26	-34	-18.5	-27.7	3.3	-55	-101	45.6
381	2.5	-19	-31	-18.5	<-30	8.2	-55.5	-94	46.0
383	2.5	-25	-33	-25	<-30	3.5	-52.5	-94	52.6
386	2.4	-22	-33	-20	<-30	4.7	-55	-97	50.1
393	2.6	-36	-38	-19.3	<-30	2.7	-52.5	-96	37.1
475	+	+	+	-22	<-30	7.5	-50.5	-103	34.0
477	+	+	+	-23	<-30	3.0	-51.5	-97	30.3
489	2.6	-36	-36	-27	<-30	1.5	-57.0	-108	41.9
490	2.6	-36	-36	-23	<-30	1.5	-56.5	-100	42.9

*   Theoretical Value (Hamming Weighting); 2.6 nsec.

** Theoretical maximum sidelobe level (Hamming Weighting); -42.8 dB

+   Not Available

†† Temperature at which the RAC chirp slope matches the design slope of 10/512 microsec per MHz

Simulation experiments on the system response to two closely spaced targets were performed, using RAC 383 in the ARWBC LC channel. The simulated target pair was generated from the input ALCOR WB ramp by dividing the signal into two paths, introducing a differential delay by adding a length of coaxial cable to one path, and then recombining the two paths. Figure 10 shows IF log amplifier output, using a 31 inch cable length to produce a 3.8 nanosecond target separation. In Figures 10a, 10b, and 10c the trailing signal relative amplitude is 0, -10 and -20 dB respectively. In Figures 10d and 10e the leading target is attenuated by -10 and -20 dB respectively. The two targets are just resolved in Figure 10a. The variations in apparent signal width with relative signal strength is attributable to interference between the two targets and to a slight variation in the delay due to the electrical length of the pads used to achieve the relative attenuation.

AVAILABLE DYNAMIC RANGE

The available signal dynamic range of the RAC is, in principle, limited at one extreme by the non-linearities in the surface wave propagation and at the other extreme by the thermal noise level. In practice, in the devices considered here, the maximum input power is set by breakdown of the input transducer so that the device cannot be operated safely in the non-linear region. Experimental evidence indicates that the input power threshold for breakdown to occur is somewhere in the range +16 dBm to +26 dBm.

Figure 11 lists the parameters that dictate the dynamic range performance of the ARWBC receiver. The driver amplifier saturation level must be at least 10 dB greater than the maximum expected signal in order to ensure amplifier linearity over the entire range of signal strengths. For the amplifier in question the saturation level is +32 dBm. In addition, if the driver amplifier is inadvertently driven into saturation, due either to a radar return from a nearby target or to arbitrary interference, the resulting input to the RAC must not exceed the maximum safe value (+16 dBm). This requires a total of 16 dB of padding between the driver amplifier and the RAC.

The ARWBC receiver gain is distributed in such a way that the system thermal noise after passing through the RAC is about 10 dB greater than the post-amplifier noise.

With the parameters shown in Figure 11 the available system dynamic range is approximately 54 dB. About 4 dB improvement is obtainable with Units 475 and 477 due to their lower insertion loss.

Figure 12 shows the deviation from linearity versus input acoustic power density for surface waves on Y-Z lithium niobate. The corresponding input power to the medium is also shown, assuming a transducer conversion loss of 16 dB. It is seen that even at the maximum permitted input power to the RAC of 16 dBm the acoustic power level is below the non-linear regime.

POWER HANDLING CAPABILITY OF RAC DEVICES

The maximum permissible power that may be applied to the RAC devices has not been established quantitatively, since an insufficient number of devices have been fabricated to permit power testing under the conditions that pertain to the present application. However, several thousand hours of reliable operation over almost four years have been accumulated with RAC devices under the conditions indicated in Figure 11.

Experiments to increase the dynamic range by reducing the padding ahead of the RAC by 10 dB resulted in catastrophic failure of two devices, as exhibited by a large increase in insertion loss. Microscopic examination of the defective units revealed extensive damage to the input transducer due to excess power.

The requirements dictated by the ALCOR WB waveform have resulted in a device that is sensitive to power in a manner that is not characteristic of most surface wave devices. Generally, SAW devices can be driven well into the non-linear region without any adverse effects. In order to achieve the operating bandwidth at the high center frequency of 1 GHz the interdigital transducer design utilizes very few electrodes (five) and an electrode spacing of only 0.8 micron, causing the transducer to break down at a power level below the value where the transition to non-linear operation occurs.

The current densities on the transducer electrodes are extremely high. Assuming that transducer failure occurred under conditions of driver amplifier saturation, with 10 dB of padding removed from the RAC input the incident power level was +26 dBm. Making the further assumption that the transducer input impedance is capacitive, the typical capacitance being 2 pf, the corresponding peak current density at the hottest spot on the transducer electrode was on the order of 10^8 amps/cm^2. No other RAC failures occurred after restoring sufficient padding to ensure that the incident power at the RAC could not exceed +16 dBm in the event of driver amplifier saturation.

RELIABILITY

The statistical sample of RAC devices is far too small for any reliability analysis. However, the evidence is that the RAC has good reliability when operated at conservative power levels. Two devices developed intermittent connections on the input microstrip line and unrelated to the RAC itself. These units operated normally after resoldering.

The initial batch of RAC's, unit No's 340, 381, and 383, were furnished in unsealed packages. Later devices were supplied in sealed housings that had been evacuated and back-filled with dry nitrogen. No deterioration in RAC performance with time could be attributed to the unsealed configuration.

CONCLUSIONS

Reflective Array Compressors have operated successfully in the ALCOR radar as matched filters to the high resolution waveform, providing a dual polarization all range detection capability for targets appearing at any arbitrary range within the pulse repetition interval and with range separations between targets as small as 4 nanoseconds. After initial installation into the system, the RAC's themselves require no additional calibration or realignment.

Installation of RAC devices in the ALCOR All Range Wideband Channel has brought about a very significant savings in equipment and the flexibility of simultaneous dual polarization operation. In addition, the matched filter has enabled an all-range wideband display to be constructed that has proven to be extremely useful as an acquisition aid for the ALCOR WB tracker.

ACKNOWLEDGMENTS

This paper represents the culmination of the work of many individuals. The RAC devices were developed in the Surface Wave Technology Group at Lincoln Laboratory under the direction of Ernest Stern. Dr's R. C. Williamson, H. I. Smith, and V. S. Dolat are primarily responsible for bringing the RAC to fruition. Their numerous innovations have jointly contributed to the excellent performance of the devices made available to the ALCOR staff.

B. Loesch designed and packaged the analog hardware which, together with the RAC components, constitutes the two matched receivers of the ALCOR All Range Wideband Channel.

We wish to thank Dr's R. C. Williamson and V. S. Dolat for supplying experimental data and photographs.

Lastly, it is a pleasure to acknowledge the contributions of members of the ALCOR staff, both past and present.

<div align="center">REFERENCES</div>

1. Camp, W. W., et al, 1971, "ALCOR - A High Sensitivity Radar with 0.5 Meter Range Resolution", IEEE International Convention Digest, pp. 112-113.

2. Williamson, R. C. and Smith, H. I., April 1973, "The Use of Surface-Elastic-Wave Reflection Gratings in Large Time-Bandwidth Pulse-Compression Filters", IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-21 No. 4, pp. 195-205.

3. Williamson, R. C., Dolat, V. S. and Smith, H. I., 1973, "L-Band Reflective Array Compressor with a Compression Ratio of 5120", IEEE Ultrasonics Symposium Proceedings, pp. 498-499.

4. Williamson, R. C., May 1976, "Properties and Applications of Reflective-Array Devices", Proc. IEEE Vol. 64, No. 5.

5. Williamson, R. C., 1974, "Problems Encountered in High-Frequency Surface Wave Devices", IEEE Ultrasonics Symposium Proceedings, pp. 321-327.

6. Klauder, J. R., Price, A. C., Darlington, S., and Albersheim, July 1960, "The Theory and Design of Chirp Radars", Bell Syst. Tech. J. Vol. 39, pp. 745-808.

7. Martin, T. A., 1976, "Low Sidelobe IMCON Pulse Compression", IEEE Ultrasonics Symposium Proceedings, pp. 411-414.
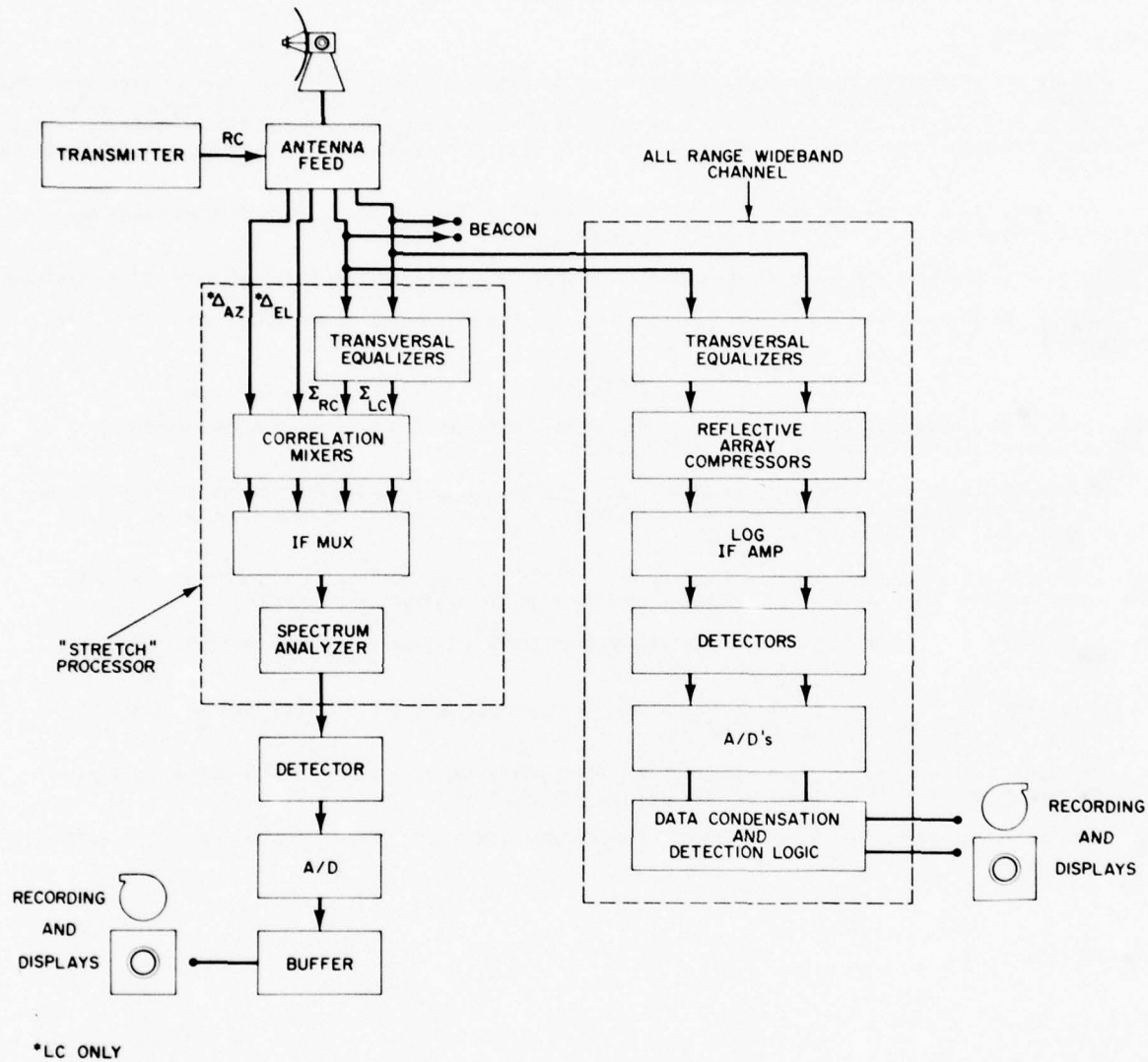
Fig. 1   ALCOR receiver configuration.

- MISSION ALIGNMENT REQUIRED
- ENVIRONMENT CRITICAL

- FIXED ALIGNMENT
- INHERENTLY STABLE
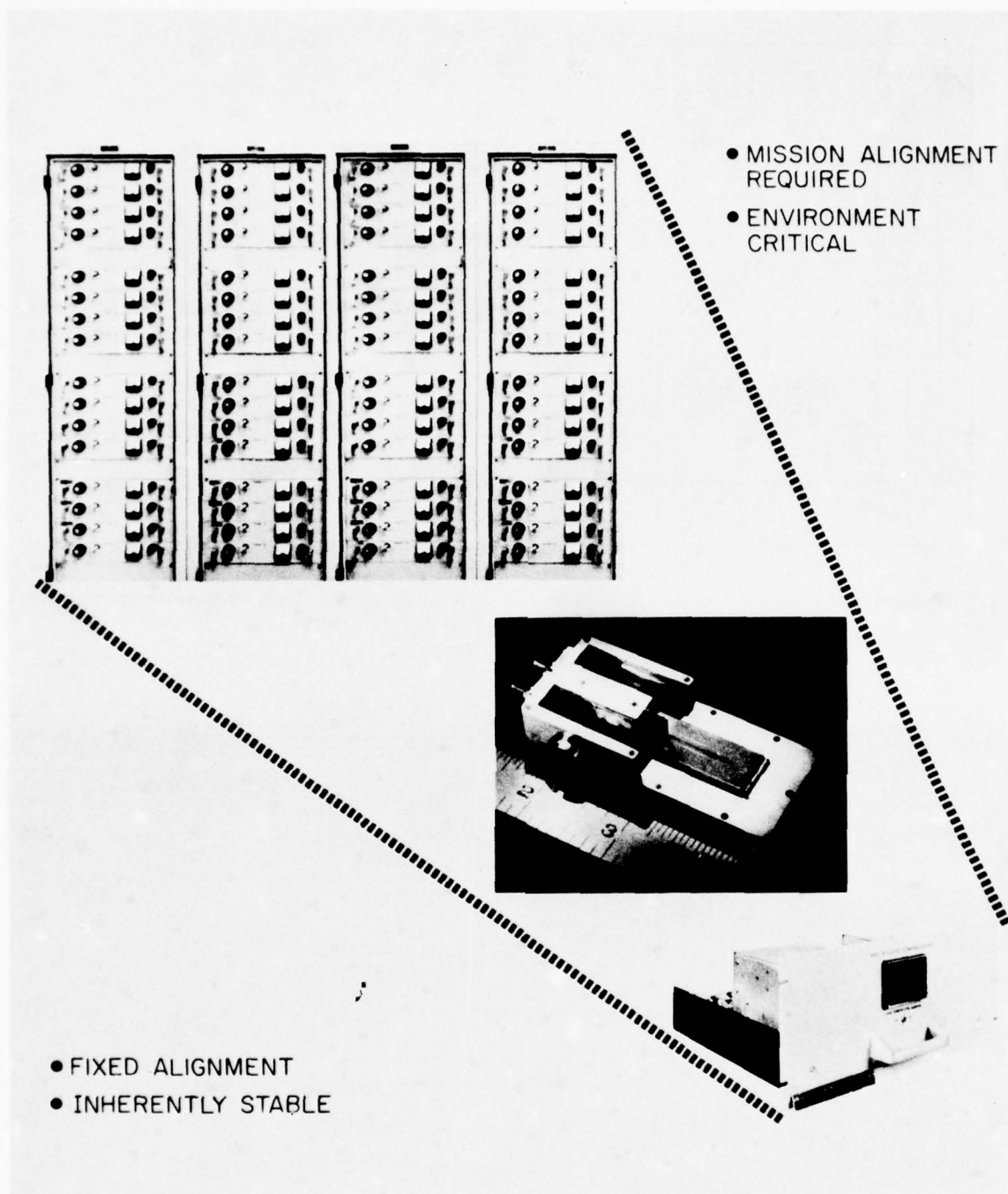
Fig. 2   Illustration of equivalent receiver hardware for the ALCOR all range wideband channel using either reflective array compressors or the band-partitioned receiver approach. Insert shows (a) the RAC assembly for a single receiver channel, (b) two L-band RAC assemblies in separate thermostatted ovens for a two channel, dual polarization receiver, (c) two rack-mounted solid state band-partitioned receivers, each comprising 32 subchannels, operating at an IF frequency of 40 MHz.

$$\text{DISPERSIVE AREA} = N\left(\frac{T_0 W_0}{2}\right) = \frac{TW}{2N}$$

$$\text{CHANNEL BANDWIDTH } W_0 = \frac{W}{N}$$

NARROWBAND, NONDISPERSIVE DELAY LINES
PROVIDE BULK OF REQUIRE AREA (shaded)
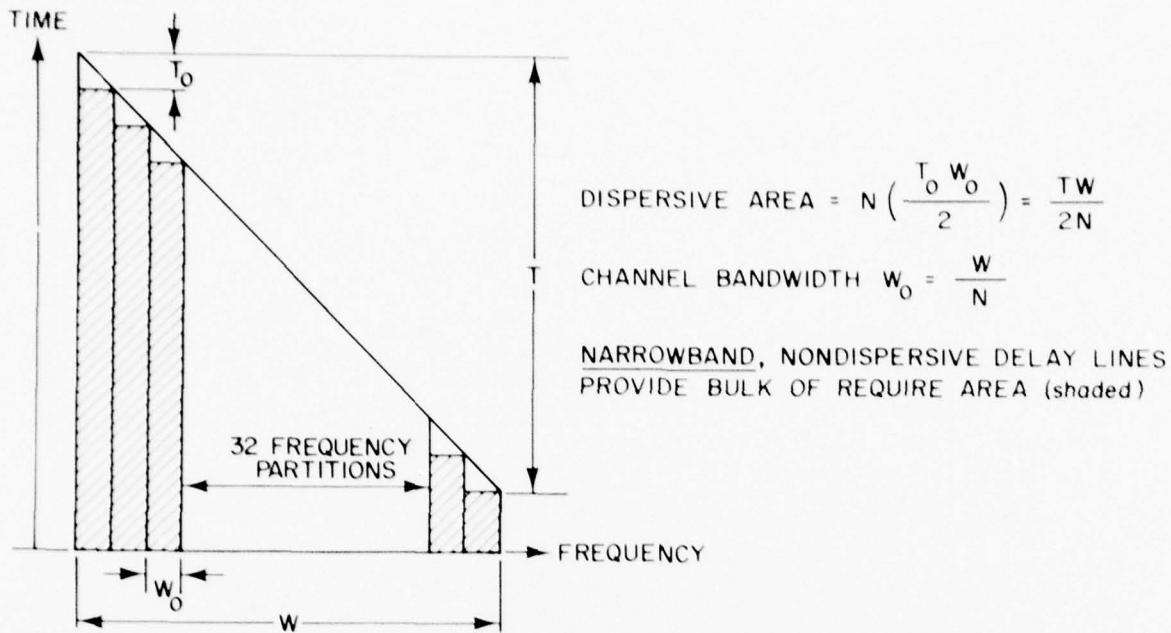
Fig. 3 Principle of operation of the all range processor band partitioned receiver. Each subchannel requires both a dispersive and a non-dispersive component of delay.
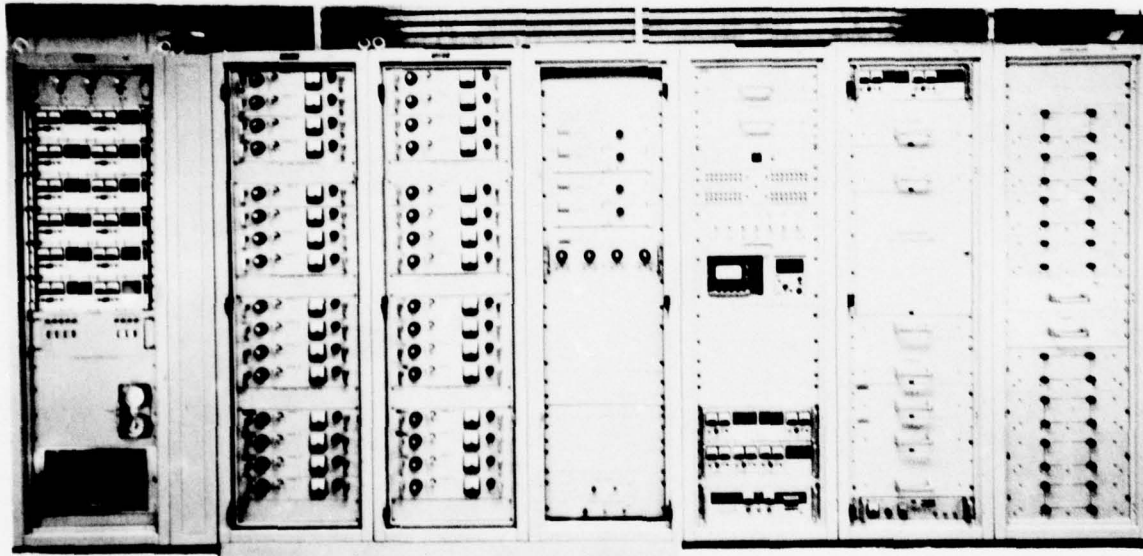
Fig. 4 The ALCOR all range processor. The total rack space contains a single band-partitioned receiver for processing a single selectable polarization, coherent mixers, power supplies, fixed delays, and transversal equalizers Dual channel, RAC assembly replaces six of the seven equipment cabinets, but the IF's are retained.
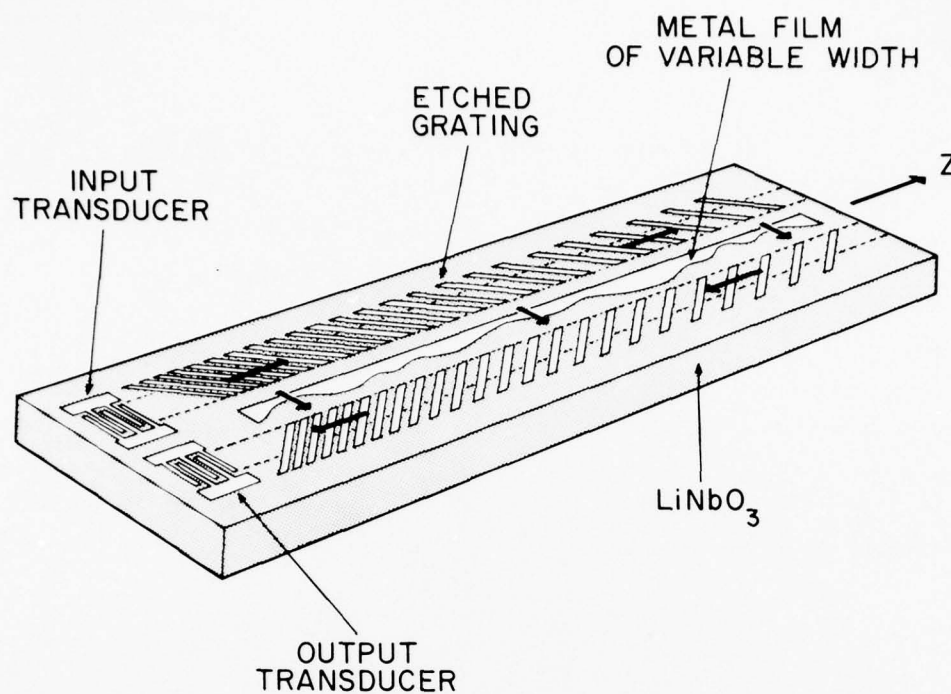
Fig.5   Schematic of reflective array compressor on LiNbO₃ substrate. The metal film of variable width introduces an additional delay of the surface wave to provide phase compensation.
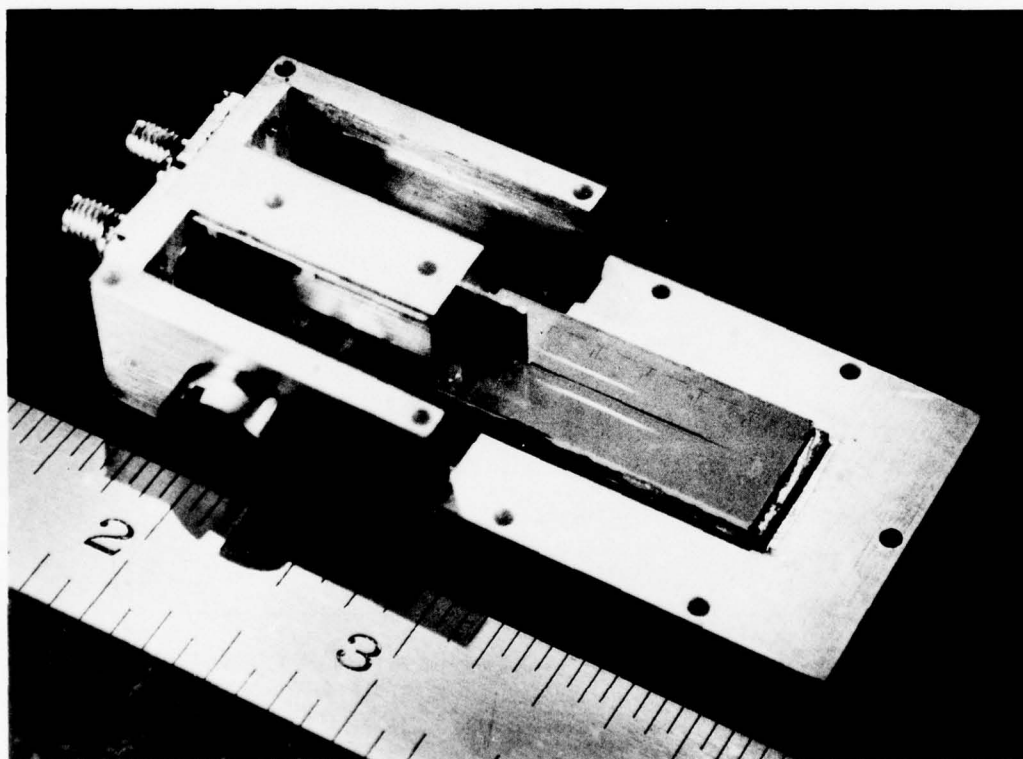
Fig.6   Detail of RAC assembly. Portions of the reflection grating on the substrate are visible.
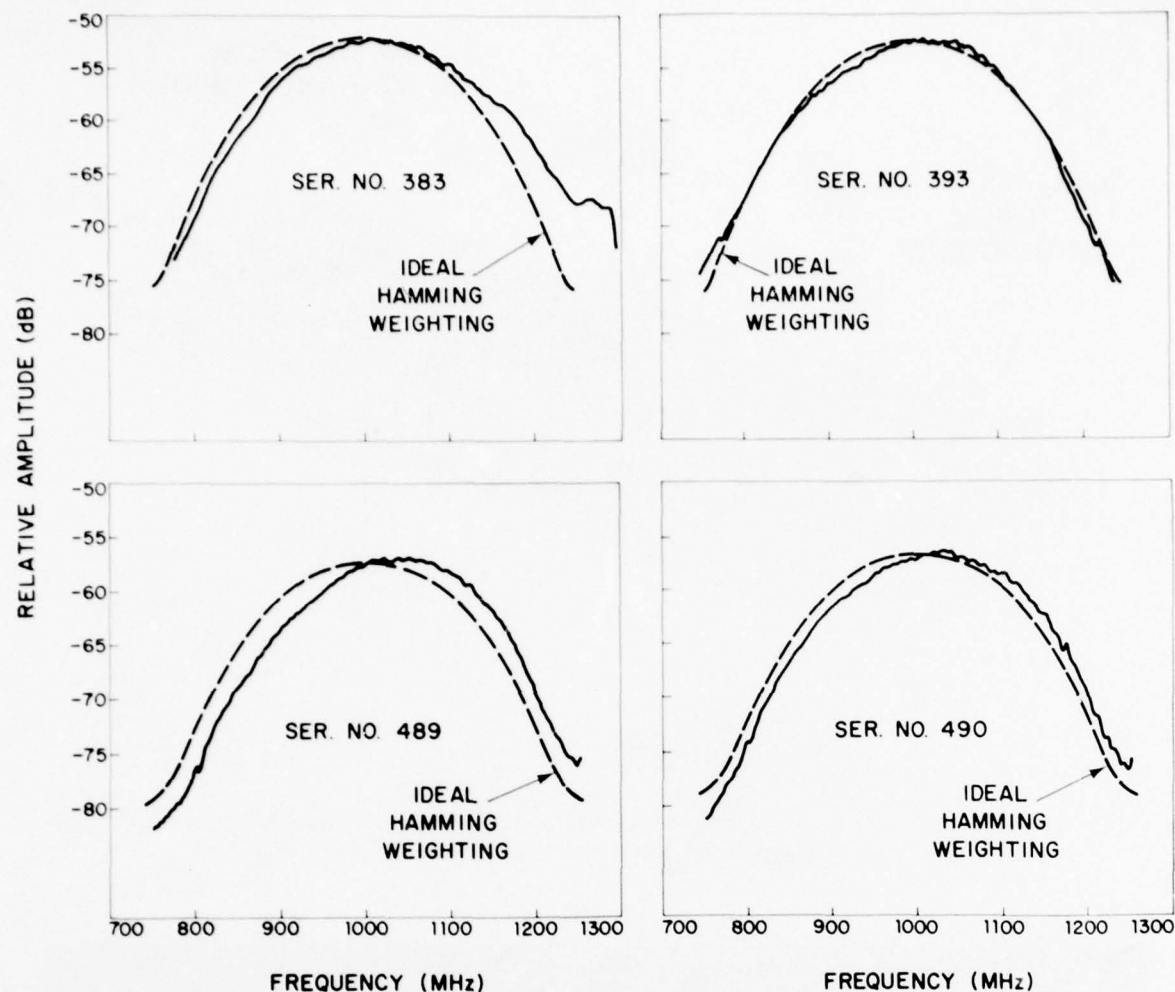
Fig.7(a)   Amplitude response for four reflective array pulse compressors.
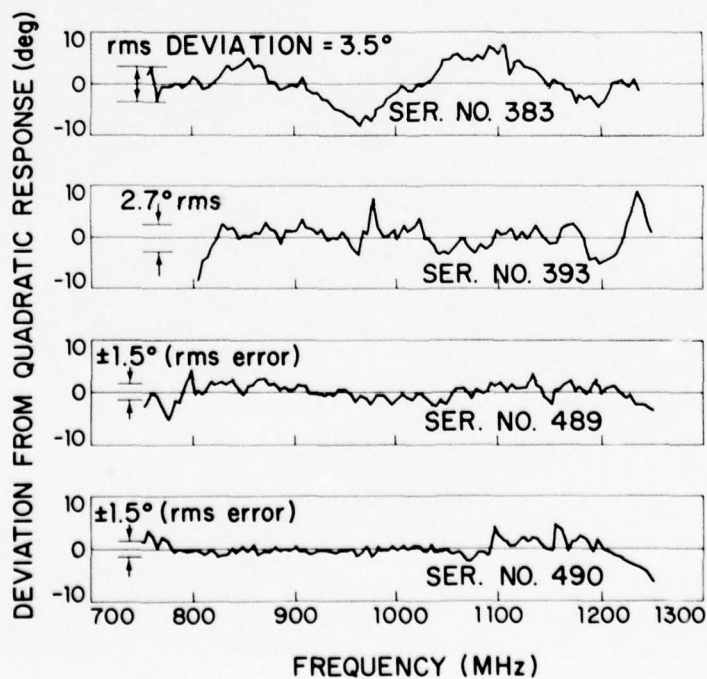
Fig.7(b)   Compensated phase response for four reflective array pulse compressors.
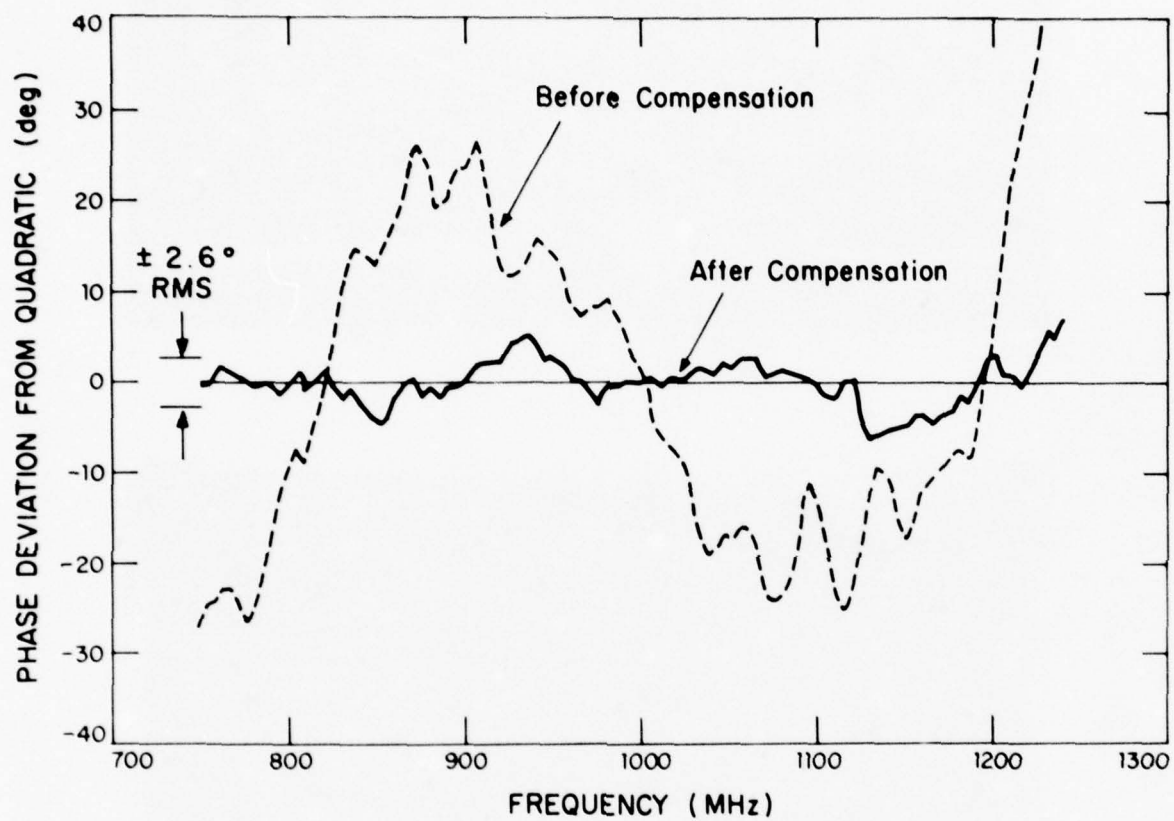
Fig.8 Typical uncompensated and compensated phase responses for a reflective array compressor.
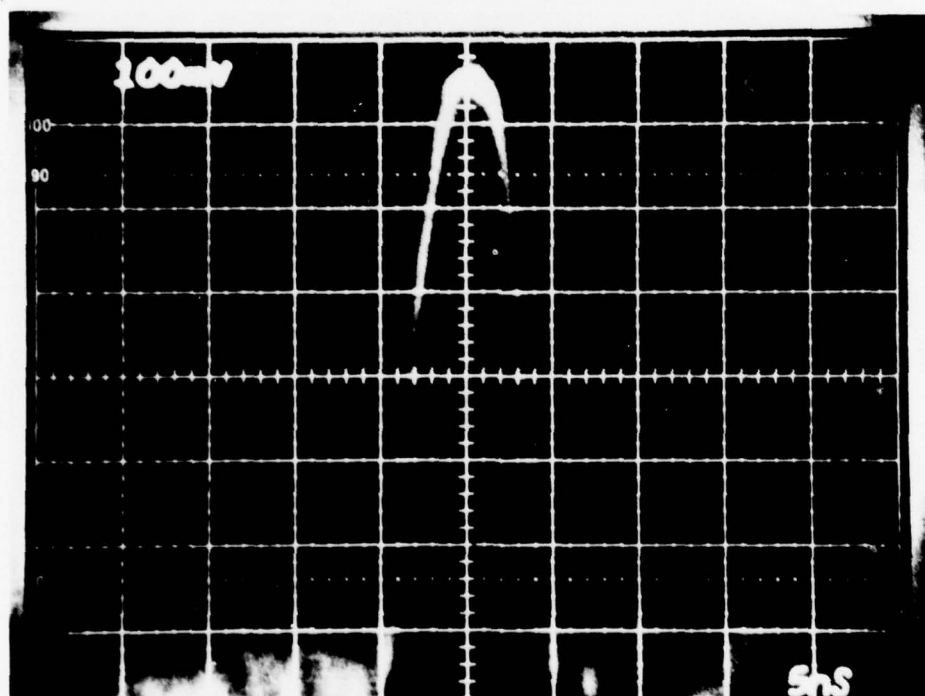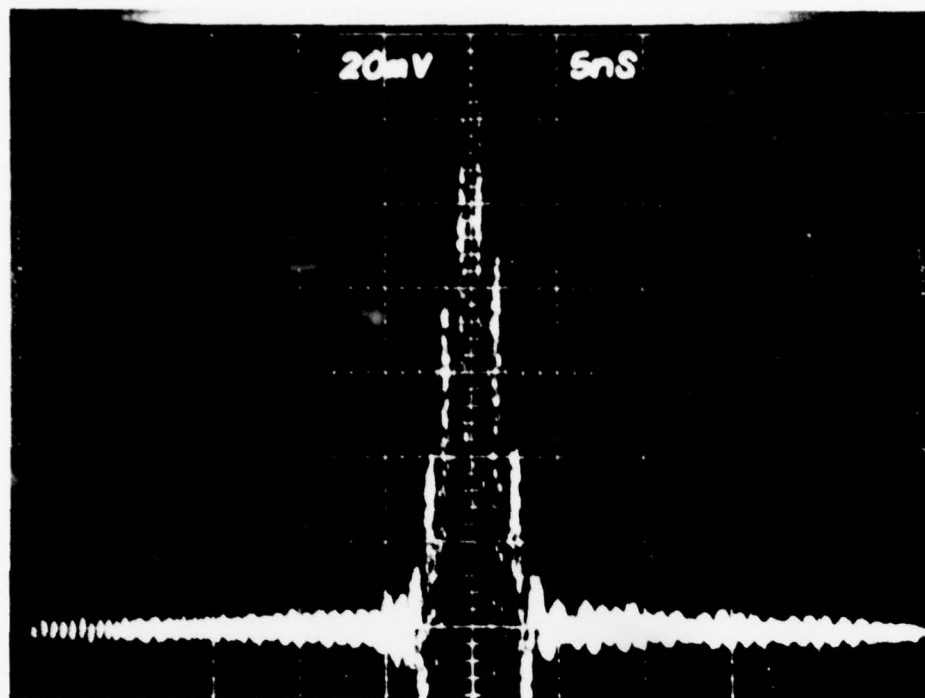
Fig. 9   Compressed pulse response of RAC serial No. 383 in ALCOR all-range wideband channel.
(a)   Carrier response at output of the RAC.
(b)   Detected pulse at output of the log IF amplifier.
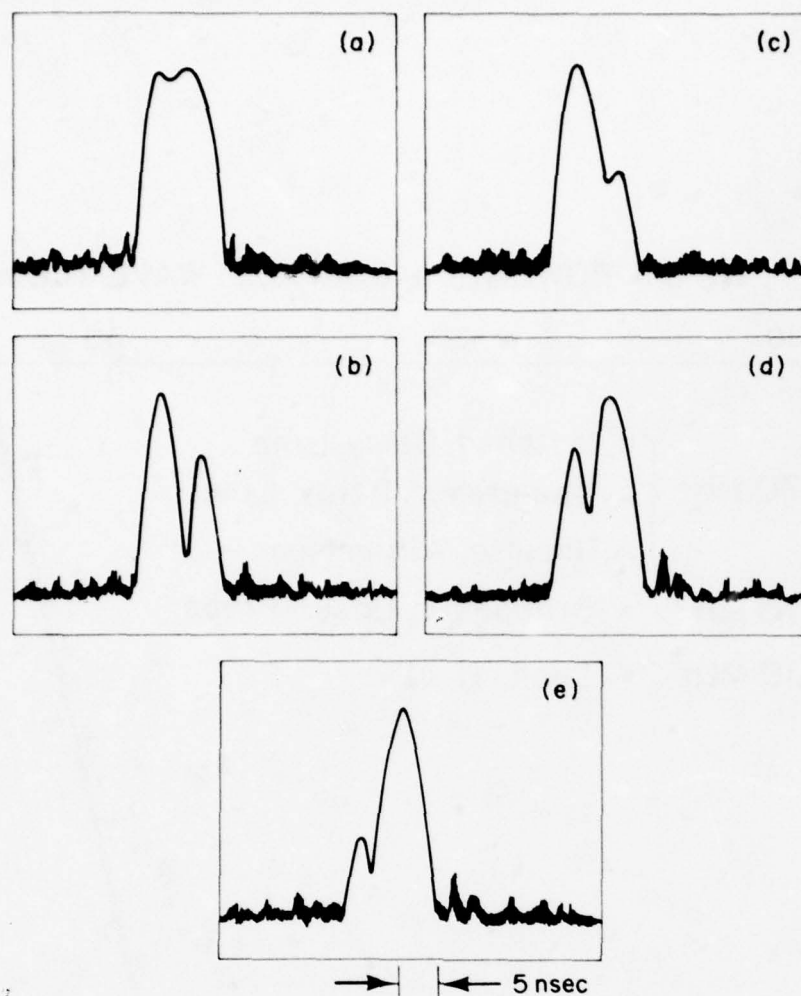
Fig.10 ARWBC response to a simulated pair of closely spaced targets. Cable delay between targets is 3.8 nano-seconds. In (a), (b), and (c), the relative trailing target amplitude is 0, − 10, and − 20 dB respectively. In (d) and (e) the relative leading target amplitude is − 10 and − 20 dB respectively.

DRIVER MAXIMUM LINEAR OUTPUT	22.0 dBm
TOTAL PADDING	−17.0 dB
RAC AVERAGE CW LOSS	−54.3 dB
PULSE COMPRESSION GAIN	30.0 dB
SYSTEM NOISE LEVEL	−73.4 dBm
AVAILABLE DYNAMIC RANGE	54.1 dB
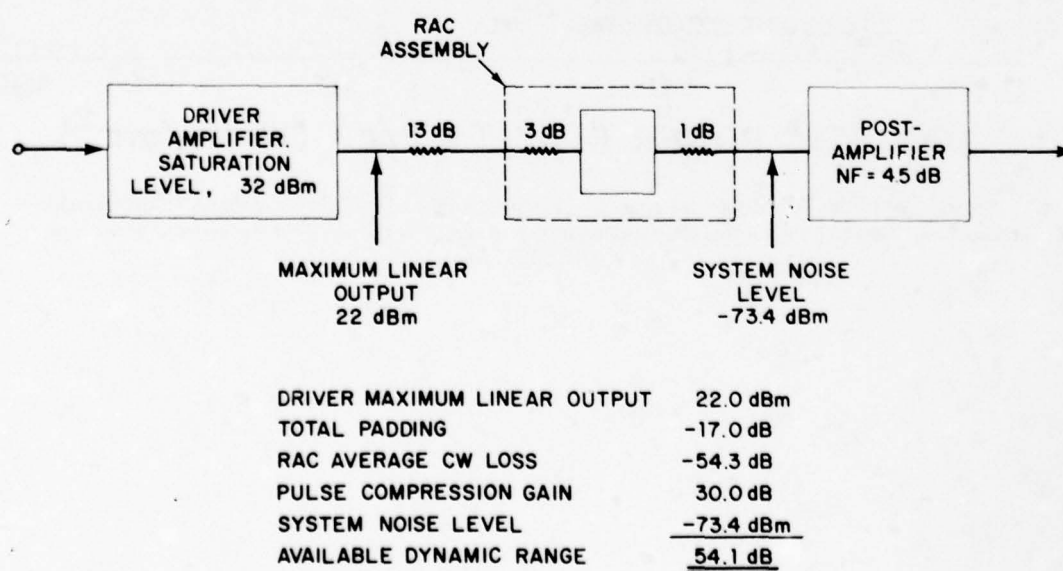
Fig.11 Dynamic range parameters for reflective array compressor in ALCOR all-range wideband channel.
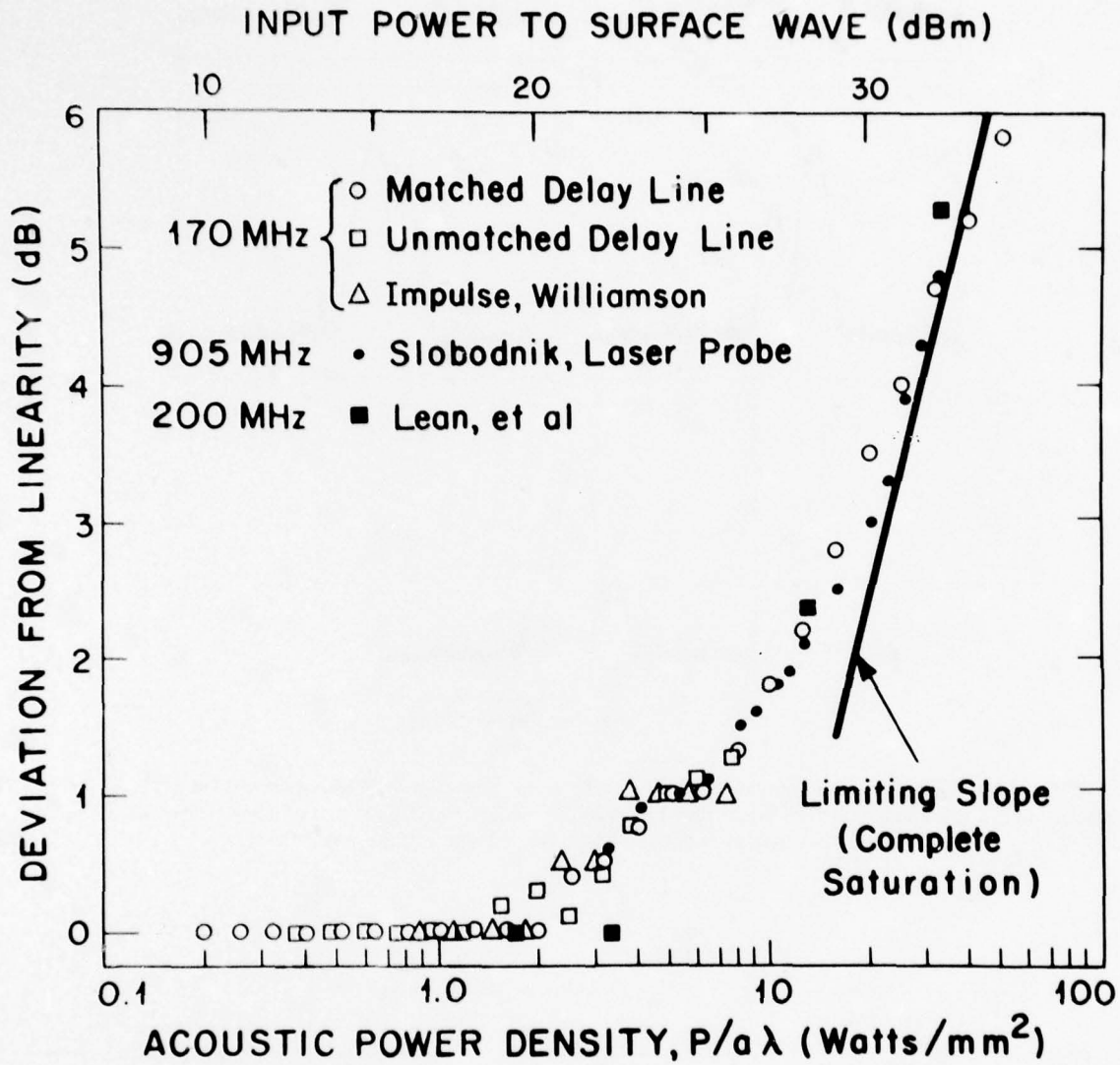
Fig. 12 Acoustic saturation for surface waves on Y-Z lithium niobate. Input power assumes 16 dB transducer conversion loss. P is the power in the propagating surface wave, 'a' is the width of the acoustic beam, and $\lambda = 3.5\mu$.

A NOVEL SIGNAL INTEGRATOR USING CCDs

C.P. Traynar, P.C.T. Roberts, R.G. Taylor,* J.D.E. Beynon, G.G. Bloodworth,
University of Southampton,
Southampton,
England.

* R.G. Taylor is with Admiralty Surface Weapons Establishment, Portsmouth, England.

SUMMARY

The implementation of two types of CCD signal integrator using non-recursive and recursive integration is described. In order to minimize the effects of transfer inefficiency, a parallel transfer scheme is employed in which information is sequentially gated into storage areas with an on-chip tapped CCD register. Test devices are described that were used to investigate the proposed implementations and, in particular, the on-chip addressing scheme. The results presented here demonstrate the feasibility of the techniques and also show good agreement with the predicted performance.

The data obtained from the test devices enabled a modified integrator to be designed, incorporating several improvements which are also described.

1. INTRODUCTION

The usefulness of recursive signal integrators to obtain signal-to-noise improvements in periodic signals, in particular radar returns, has long been recognized. Charge transfer devices (CTD) have recently been used (Roberts et al, 1974) to implement the delay element of such systems since they offer considerable advantages over alternative techniques which have tended to be bulky, expensive and of limited versatility. Unfortunately, however, the signal/noise improvement which can be obtained using an integrator employing a CTD in the "conventional" (i.e. serial) way is limited by the build-up of signal residuals caused by transfer inefficiency (Chowaniec et al, 1975). For example, a recursive CCD integrator of 100 bits designed to sum an effective sample size of 10 exhibits a signal-to-residual level of $\simeq 35dB$ if the transfer inefficiency is 2×10^{-5} per bit, a state-of-the-art figure for a buried channel CCD. In a serial system, however, most of the transfers that the charge packet undergoes do not contribute to the signal processing function as such. If these "redundant" transfers could be eliminated then a considerable improvement in performance could be achieved; e.g. for the example quoted above the signal-to-residual ratio would become 75dB. We have designed a novel scheme in order to realize this better performance in which a parallel transfer CCD implementation is adopted (Traynar et al, 1977a).* Two basic types of radar video integrators are described in this paper; one uses non-recursive integration in which the last 'n' radar returns are summed with equal weighting, the other uses recursive integration in which the radar returns are summed with exponential weighting.

The operation of the non-recursive integrator is illustrated in figure 1(a). Each radar range bin is implemented in the form of a tapped 'm' bit CCD terminated with a drain diffusion. Successive returns are loaded into the CCDs by sequentially operating the analogue gates G_1, G_2, G_3 ... G_N. Thus, at any instant, each CCD contains the last m echoes from one particular range bin. Since each CCD has all its taps connected together, integration over the m samples is achieved; this signal is then gated to the output through gates G_1', G_2', G_3' ... G_N'. Although the number of radar returns summed in this scheme is essentially fixed by the chip artwork, a limited degree of choice could be provided by segmenting the summing bus bars with MOST couplings. In the recursive system (illustrated in figure 1(b)), the returns are similarly gated sequentially into each range bin storage well. Summation in this case, however, is achieved in the manner used in conventional recursive integrators except that each range bin is processed individually. This technique allows an arbitrarily large number of returns to be summed (albeit with exponential weighting), and under certain circumstances enables a larger signal-to-noise improvement to be achieved than is possible with the non-recursive system.

In both integrators each radar return undergoes a minimal number of CCD transfers (two per processing step in a 2-phase system) and so residuals caused by transfer inefficiency are maintained at a very low level. The parallel input and output organization also allows the number of range bins in the total system to be increased as desired, without degrading the performance, simply by cascading the requisite number of devices together.

It will be readily appreciated that, in order to realize the systems described above, an on-chip addressing technique is essential to minimize the number of external connections. We have implemented such a system by means of a "range gate register" (RGR), which comprises a floating-gate-tapped, two-phase CCD. This device is operated as a digital shift register and is loaded with a single full charge packet to represent a '1'. As the charge packet is shifted past each tapping point, an address pulse is generated which operates the analogue gates depicted in figures 1(a) and 1(b). In order that chip complexity is not increased by the inclusion of MOS or resistive biassing for each of the individual floating-gates

---

* Although a parallel architecture has been described by Tiemann et al (1974), they used MOS shift registers designed to operate at up to 4MHz to store information samples sequentially in an array of surface charge transistors (SCT) storage sites. The use of CCD structures for the switching and storage elements has the potential advantage of higher operating speeds and less chip complexity for the same system function.

(MacLennan et al, 1975), a novel scheme is employed in which each gate is allowed to float completely, any loading being purely capacitive. As a result, the floating-gates acquire their bias voltage by virtue of the lateral leakage of charge along the oxide from the adjacent switched electrodes. Unlike capacitive biassing structures fabricated with multilevel processes (Wen, 1974), capacitive loading is minimized and so large floating-gate potential swings can be achieved which can drive the analogue gates directly without intermediate amplification.

2.   THE NON-RECURSIVE SCHEME - TEST CHIP

Several implementations of the non-recursive integrator have been considered, using either voltage or current tapping methods in conjunction with external or on-chip clocking of the range bin CCDs; the combinations which minimize chip size and complexity appear to be those which employ

(i)   current sensing of the range bin CCDs with a clocking pulse derived from the RGR, and

(ii)   voltage sensing of the range bin CCDs and off-chip clocking triggered by the address bit arriving at the end of the RGR.

Since both the sensing methods considered involve non-linearities due to depletion capacitance, it was decided to fabricate a simple test chip in the form of a 50 bit two-phase CCD in which these non-linearities could be investigated. In addition, the feasibility of using the simplified floating-gate biassing scheme discussed in the last section for the range bin CCDs could also be determined. The device was designed to have total gate lengths of 14μm, consisting of a 4μm thick oxide and a 10μm thin oxide part. In order that the floating-gate voltage sensing mode could be investigated thoroughly, the floating electrodes were connected directly to an on-chip MOST with an external MOST load; this not only minimized the capacitive loading on the floating electrodes but also allowed this parameter to be accurately determined for subsequent computer analysis. A photomicrograph of the complete device is shown in figure 2.

Since very little work has been done on the effects of depletion capacitance induced distortion in floating-gate sensing structures, it was decided to carry out a detailed analysis of this effect using a charge balance model. This work has been described in detail elsewhere (Traynar et al, 1977b) and only the final result will be quoted here. For a floating-gate structure containing N identical electrodes of area A under which signal charges Q_{S1}, Q_{S2}, Q_{S3} ... Q_{SN} are simultaneously injected, the final floating-gate potential $V_g(Q)$ may be evaluated using the following equation

$$\sum_{1}^{N} A Q_{Sn} = \sum_{1}^{N} A C_{OX} \left[V_o^2 + 2V_o V_g(0) \right]^{\frac{1}{2}} - \sum_{1}^{N} A C_{OX} \left[V_o^2 + 2V_o \left(V_g(Q) - \frac{Q_{Sn}}{C_{OX}} \right) \right]^{\frac{1}{2}}$$

$$+ \left[V_g(0) - V_g(Q) \right] C_S \qquad \qquad ... 1$$

where $V_g(0)$ is the initial floating-gate potential and

$$V_o = \frac{N_d\, q\, \varepsilon_s}{C_{OX}^2}$$

where N_d is the substrate doping density

q is the electronic charge

ε_s is the dielectric constant of silicon

C_{OX} is the gate oxide capacitance per unit area

and C_s is the constant stray loading capacitance.

For N > 1, this equation must be solved iteratively to obtain $V_g(Q)$ and we have developed a computer program in which all types of CCD structure (including split-weight tapped, transversal filters) may be evaluated. For the stepped oxide two-phase structure under consideration in this paper, equation I may be easily modified to include the loading effects of the thick oxide portion of each phase gate simply by including more "gates" in the model of the structure; no charge is stored of course under the thick oxide portions of the structure - for computational purposes these parts of the electrodes merely increase the loading capacitance. It is not difficult to make reasonable estimates of all the parameters in equation 1, with the exception of the stray capacitance C_s. By using the MOS buffer amplifier mentioned previously it was possible to measure the floating-gate potential without external loading; therefore an accurate experimental determination of C_s was achieved. In this way an analysis of the test structure was carried out the results of which are shown in figure 3.

Experimental measurements were performed on the test device using an HP310A wave analyser buffered by a Brookdeal Precision AC Amplifier type 425. The linearity of this amplifier is quoted as 0.001% and it had an observed noise level of at least 70dB below the fundamental signal component with the input short-circuited. The input signal was connected to the input gate via a simple RC level shifting network. With

a lv pk-pk 1kHz signal at the input gate, the harmonic components at this node were less than 60dB below the fundamental. The distortion introduced by the gated charge detector was found to be negligible and so the linearity of the charge injection scheme could be optimised by adjusting the bias on the input gate to give the lowest harmonic content at the CCD output. This resulted in the harmonic components generated at the input being less than 50dB below the fundamental. Also, harmonics generated by the floating-gate sense amplifier were less than 60dB, and so we felt justified in interpreting the experimental results as being direct measurements of the floating-gate non-linearity. These measurements over an input signal range up to 0.05 times the clock frequency of 200kHz are also plotted in figure 3. It can be seen that the model gives a very close fit to the measured response, leading us to feel confident that we can reasonably accurately predict the linearity of a complete non-recursive system.

3. RGR AND RECURSIVE INTEGRATOR TEST CHIP

It has been indicated that successful operation of the RGR is essential to both the parallel transfer integration schemes described in this paper. In order to test this facility and its ability to interface correctly with the analogue gates, a recursive integrator test chip was designed using the analytical techniques described earlier (i.e. using equation 1), containing a 4 bit RGR and three storage sites. A schematic diagram of this device is shown in figure 4 and it can be seen that one RGR performs both the input and output gating; this feature is desirable since it obviates the need for a separate output RGR and so considerably reduces active chip area. Although it is proposed that the RGR should drive each analogue gate separately, MOS buffers were included in this prototype chip, thus enabling the low speed performance of the RGR to be evaluated without excessive loading effects. (Subsequent designs will feature exclusively CCD structures, thus enabling relatively high speed operation (10MHz) and high packing density to be achieved.)

To simplify implementation of the analogue gates, which must operate effectively at up to 10MHz data rate, each storage site and associated input and output gates are realized in the form of a single bit, two-phase CCD. This structure (shown in figure 5) is operated in the $1\frac{1}{2}$-phase mode using the same clock pulse that is used to drive the RGR. The input analogue gate is simply the input section of this CCD structure, comprising electrodes ϕ_1, ϕ_2^{j+1} and G_2. Charge injected by this method has demonstrated a high signal voltage-to-charge packet size linearity (Haken, 1975). Output gating is accomplished by pulsing ϕ_2^j and thus transferring the signal charge stored beneath G_2 to the output sense diode.

Operation of the integrator is as follows: Consider an address pulse from the jth position of the RGR and buffer arriving at the ϕ_2^j electrode. Signal charge previously stored in the G_2 well will flow onto the output diode (which is connected in parallel with all other output diodes on the chip such that the total capacitance at this node does not exceed ~5pF) (see figure 5(a)). When the next clock pulse drives the RGR signal charge representing the '1' to the next tap position, ϕ_2^j will turn off. Simultaneously, all ϕ_1 electrodes are pulsed so that their wells are deep enough to accept charge from their adjacent input diffusions (figure 5(b)). During the ON period of the clock, the signal voltage on the output diode is sampled and the diode is reset to a reference level. As the clock goes off, the '1' in the RGR is shifted along to the next bit, causing ϕ_2^{j+1} to pulse on. Since the ϕ_1 well is now in the "supply" position, the well formed beneath the gate G_1 will fill with charge proportional to the new signal voltage V_{ig} (figure 5(c)). On the next clock cycle this signal charge is isolated in the G_2 well (figure 5(d)); thus within two clock periods, a given storage site can be accessed and updated. A photomicrograph of the test chip can be seen in figure 6. To reduce chip complexity in this test device no sample/hold circuitry was included. The shadow gap technology (Baker, 1973) invented and developed here, was used to fabricate the sub-micron separations between the CCD electrodes; this also ensures a minimum gate-to-gate capacitance unlike most other structures and gives only a small clock pulse breakthrough to the ϕ_2 floating-gates. Although the device is basically a two-phase CCD structure, the three-phase (2-level metallization) shadowing technique was employed to define the split gate G_1 and G_2 of figure 5. This process has the further significant advantage that small feature dimensions can be readily achieved without special alignment procedures. In particular, the gate G_1 is only 4μm long and, together with its interconnection track, it is defined completely with the shadow technique. The aluminium was evaporated by electron-beam from a carbon crucible which, after low temperature H_2 annealing, gave very uniform turn-on voltages over a complete 2" wafer of < 100 > oriented material (typical values are 2v ± 0.2v for thin oxides of 1200 Å and 6v ± 0.5v for thick oxides of 5000 Å).

The dynamic operation of the RGR is shown in figure 7; the input pulse is shown at the top of the photograph and the four inverter outputs beneath. It can be seen that the input pulse is delayed by the appropriate number of clock cycles at each inverter output. Overlooking the small trailing charge (which will be discussed later), there is a variation in pulse amplitude due to a variation in the gains of the inverters from 1.4 to 2.7. The variation is in excess of that expected due to over etching, threshold variations, etc and is somewhat anomalous in that, on all of the devices tested, the gain decreased from right to left in the photomicrograph of figure 6. The cause of this effect has yet to be established. The small amount of trailing charge shown in figure 7 is the result of exaggerated transfer inefficiency caused by the operating mode of the RGR. If the charge packet representing the '1' is increased, the effect becomes more marked, figure 8; it can be seen to be not a true transfer inefficiency effect, since a third trailing charge does not appear. The effect is in fact due to the structure of the floating-gate tap which, it will be recalled, is simply a thick/thin oxide phase gate. The falling potential of the floating-gate in response to the charge packet filling the thin oxide storage section eventually blocks the filling process and gives rise to a bucket brigade type (i.e. incomplete) transfer mode with large amplitude trailing packets. If it is desired that the maximum size charge packet is transferred along the RGR, this effect can be mitigated by either completely collapsing the ϕ_1 wells, or by employing a split floating-gate structure as described in the next section.

The transfer response of each floating-gate was measured by transferring a string of variable sized charge packets along the RGR and subsequently characterizing the inverters. Experimental results are shown in figure 9 and, as can be seen, good agreement between the measured and theoretically predicted values exists.

These initial results clearly indicate that an on-chip addressing scheme employing a tapped CCD is feasible, and further, in view of the fact that large address pulses can be obtained, that direct connection to the analogue gates is possible.

Unfortunately, in spite of the success achieved with the operation of the RGR on those devices studied, the rest of the chip could not be operated as a recursive integrator due to parasitic transistor action occurring between the bus lines to the storage sites. However, since these untested functions employ conventional CCD techniques, rather than correct the fault we proceeded with the design of another layout which incorporated a number of improvements which are described below.

4.   THE IMPROVED RECURSIVE INTEGRATOR

The structure of the recursive integrator described in section 3 above was modified in the following ways.

(a)   The RGR tap electrodes are split using a three-phase, aluminium shadow-gap shown in the electron micrograph of figure 10, with the thick oxide section set to a fixed bias. This permits a greater fall in floating-gate potential before incomplete transfer occurs and also gives better electrostatic screening from the ϕ_1 clock pulses.

(b)   The RGR floating-gate taps are connected directly to the storage site input and output gates, as shown in figure 11. This eliminates the MOST buffer amplifiers and also eliminates the need for a ϕ_1 connection to the storage sites. The ϕ_2^j electrode in this arrangement delivers a supply of charge only once per sweep of the radar aerial, thereby removing the need for the split gate G_1 required in the previous implementation.

(c)   The RGR taps alternately feed eight storage sites distributed equally either side of the RGR. The lower set of storage sites have adjacent input and output diodes connected directly to each other, and subsequently to a reset transistor and input gate of an MOST inverter. The upper storage sites retain separated input and sense diodes but, in this instance, the output gate is biased by direct connection to the output diode. A combination of these two configurations could greatly increase packing density if they were adopted in a final design; i.e. the input and output diodes could be combined as a single diffusion and the need for two bus lines to bias the input diodes and output gates would be eliminated. With two sets of storage sites as described above, the compatibility of these various implementations can be evaluated.

(d)   Because of the operation of the storage sites directly from the RGR floating-gates, it is unnecessary to have a sample/hold circuit in the feedback loop to the analogue input gates. Instead, the output of charge from the jth storage site occurs with the same ϕ_2^j RGR pulse that is used to prime the input gate of that site. This means that updating of the charge stored is achieved within one cycle of the external clock voltage.

(e)   The RGR input section includes a two-phase input gate to give accurate control of the size of the '1' charge packet circulated in the RGR. The output section contains two types of detector. The first is a novel sample/hold circuit consisting of a floating-gate sense tap biased with a very high impedance load MOST and feeding a signal voltage through a triple-gate, sampling MOST switch to a simple inverter buffer. The second detector is a sense diode, connected on-chip to the final two-phase gate, to be used with either a simple resistor load or a charge sensitive amplifier. These facilities permit more detailed measurements of the performance of the RGR and easier stacking of chips with a "Refresh" function for the digital RGR word. These design improvements not only reduce the number of switching waveforms from three (ϕ_1, $\phi_{sample/hold}$, ϕ_{reset}) to two (ϕ_1, ϕ_{reset}), thereby reducing the power consumption of the driving circuitry, but also increase the packing density by a factor three. A photomicrograph of the complete chip measuring 1.25 x 1.25mm^2, is shown in figure 12(a). A closer view of the upper storage site is shown in figure 12(b) and of the lower storage site in figure 12(c).

5.   CONCLUSIONS

The work carried out so far on the devices described here demonstrates the feasibility of CCD parallel transfer structures. In addition, the operation of the RGR on the recursive test device has shown the feasibility of on-chip addressing using a tapped CCD; this approach has significant advantages over alternative techniques which are generally slower and more complex.

The RGR has proved to be quite stable under laboratory conditions; any tendency towards instability could be diminished, however, by a process modification that would increase the DC coupling between electrodes. This would reduce the effects of gate to substrate leakage and changes in ambient humidity on the floating-gate bias potential. Measurements on the second recursive integrator test device are at present in progress and results are encouraging; it is proposed that the construction of a complete decluttering subsystem will shortly be commenced using these test devices. This will allow the overall performance of the system to be evaluated for signals buried in various noise densities and to determine the effects of fixed pattern noise on the integrated signal. The applicability of this type of signal integrator (and delay line) to other systems is also being investigated.

ACKNOWLEDGEMENT

REFERENCES

BAKER, I.M., BEYNON, J.D.E. and COPELAND, M.A., 1973, "Charge-Coupled Devices with Sub-micron Electrode Separations", Electronics Letters, Vol 9, pp48-9, Feb 8.

CHOWANIEC, A. and HOBSON, G.S., 1975, "A Time Domain Analysis of CCD Video Integrators", Proc CCD '75, pp323-329, San Diego.

HAKEN, R.A., 1975, "Supply Charge Isolation - A Simple Surface Potential Equilibration Charge Injection Technique for Charge-Coupled Devices", IEEE J. of Solid State Circuits, SC-11, pp189-196.

MACLENNAN, D.J., MAVOR, J. and VANSTONE, G.F., 1975, "Techniques for Realizing Transversal Filters Using Charge-Coupled Devices", Proc IEE, Vol 112, pp615-619.

ROBERTS, J.B.G., CHESSWAS, M., and EAMES, R., 1974, "Video Integration Using Charge-Coupled Devices", Electronics Letters, Vol 10, pp169-171, May 16.

TIEMANN, J.J., ENGELER, W.E. and BAERTCH, R.D., 1974, "A Surface Charge Correlator", IEEE J. of Solid State Circuits, SC-9, No 6, pp403-409.

TIEMANN, J.J., ENGELER, W.E., BAERTCH, R.D. and BROWN, D.M., 1974, "Intra-cell Charge Transfer Structures for Signal Processing", IEEE Trans., ED-21, No 5, pp300-308.

TRAYNAR, C.P., ROBERTS, P.C.T., BLOODWORTH, G.G. and TAYLOR, R.G., 1977a, "A Novel CCD Signal Integrator", to be published in Electronics Letters.

TRAYNAR, C.P., BEYNON, J.D.E., ROBERTS, P.C.T. and PENNOCK, J., 1977b, "A Discussion of Depletion-Capacitance Induced Distortion in Surface Channel CCD Transversal Filters", IEE J. of Solid State and Electron Devices, Vol 1, No 3, pp73-80.

WEN, D.D., 1974, "Design and Operation of a Floating-Gate Amplifier", IEEE J. of Solid State Circuits, SC-9, pp410-414.
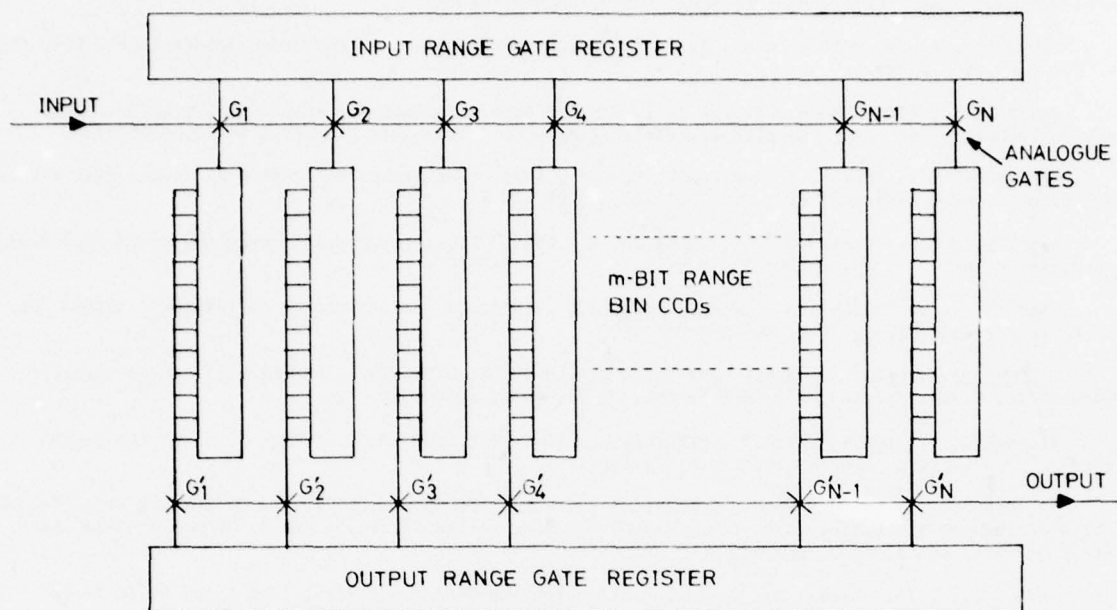
Figure 1(a)   Parallel implementation of the non-recursive integrator.
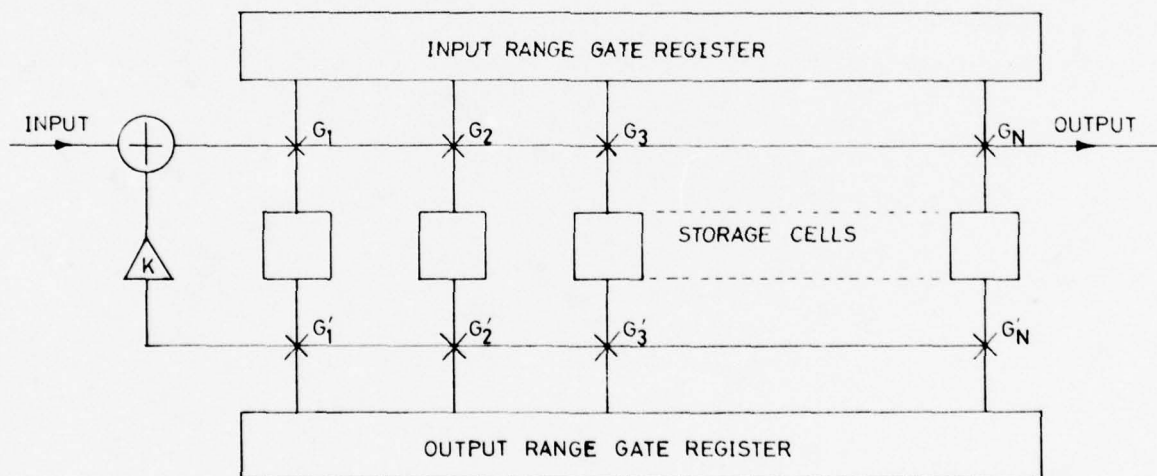
Figure 1(b)   Parallel implementation of the recursive integrator.
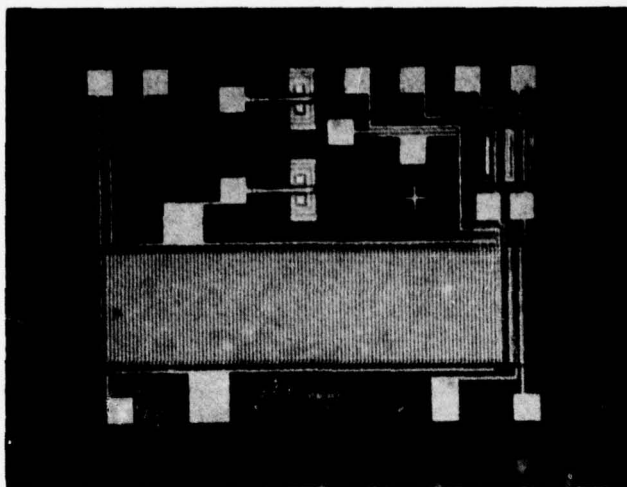
Figure 2   A photomicrograph of the two-phase CCD used to measure the linearity of the floating-gate summing technique.
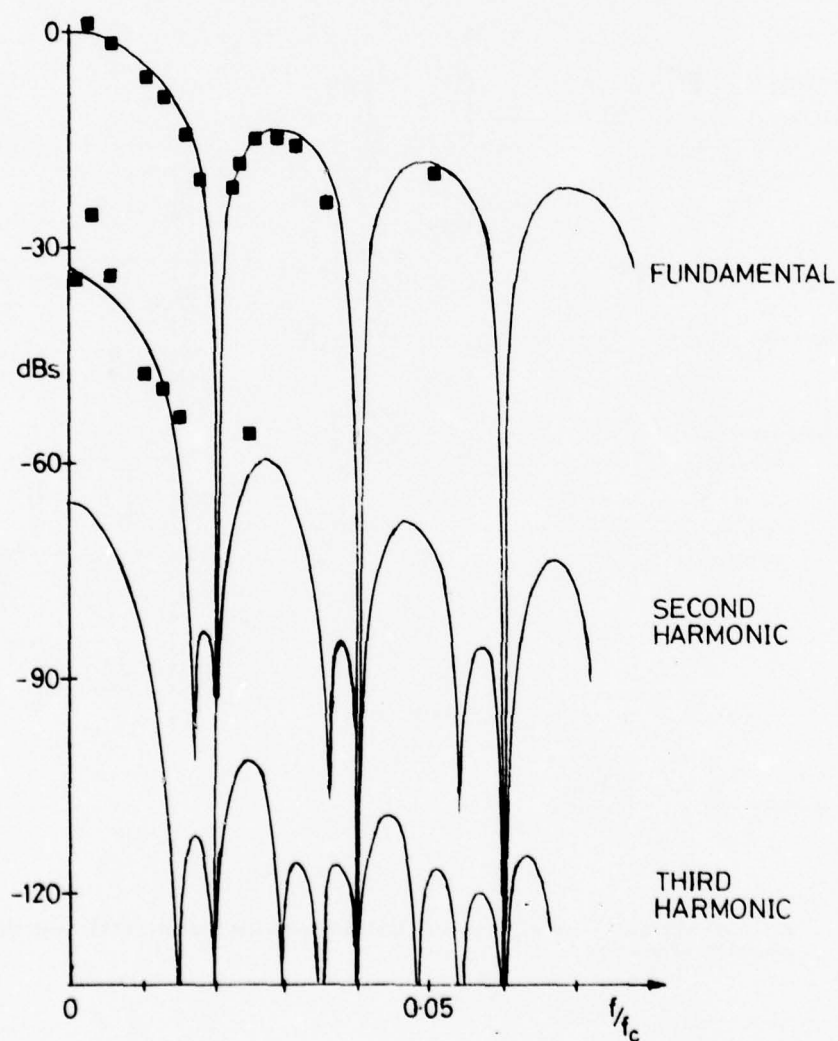
Figure 3   The harmonic distortion produced by the floating-gate summing technique. The predicted response is shown as the continuous line, with the squares representing the measured
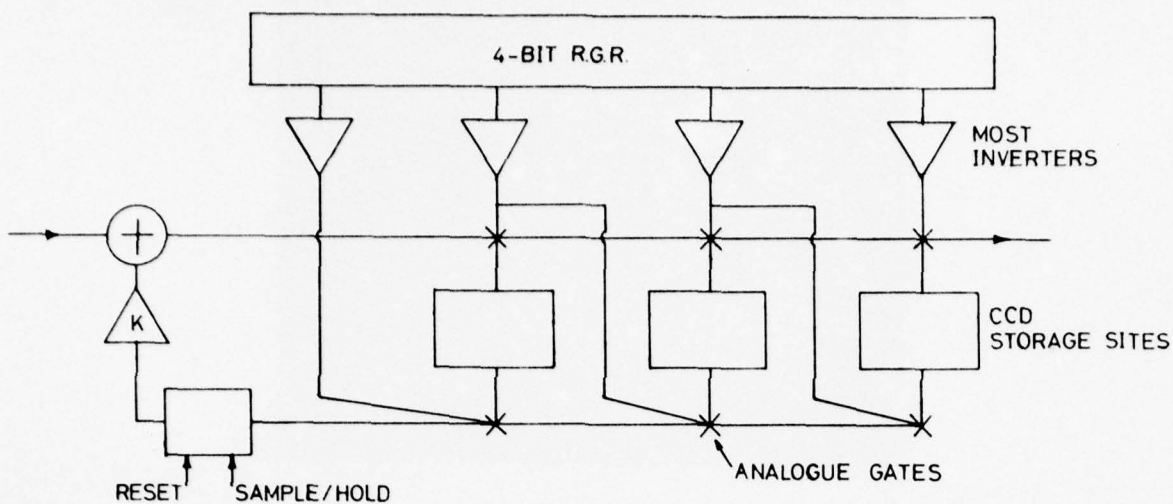
Figure 4   A schematic diagram of the chip used to investigate the performance of the RGR and its ability to interface correctly with the storage sites. The sample/hold, scaling and summing operations are carried out off-chip.
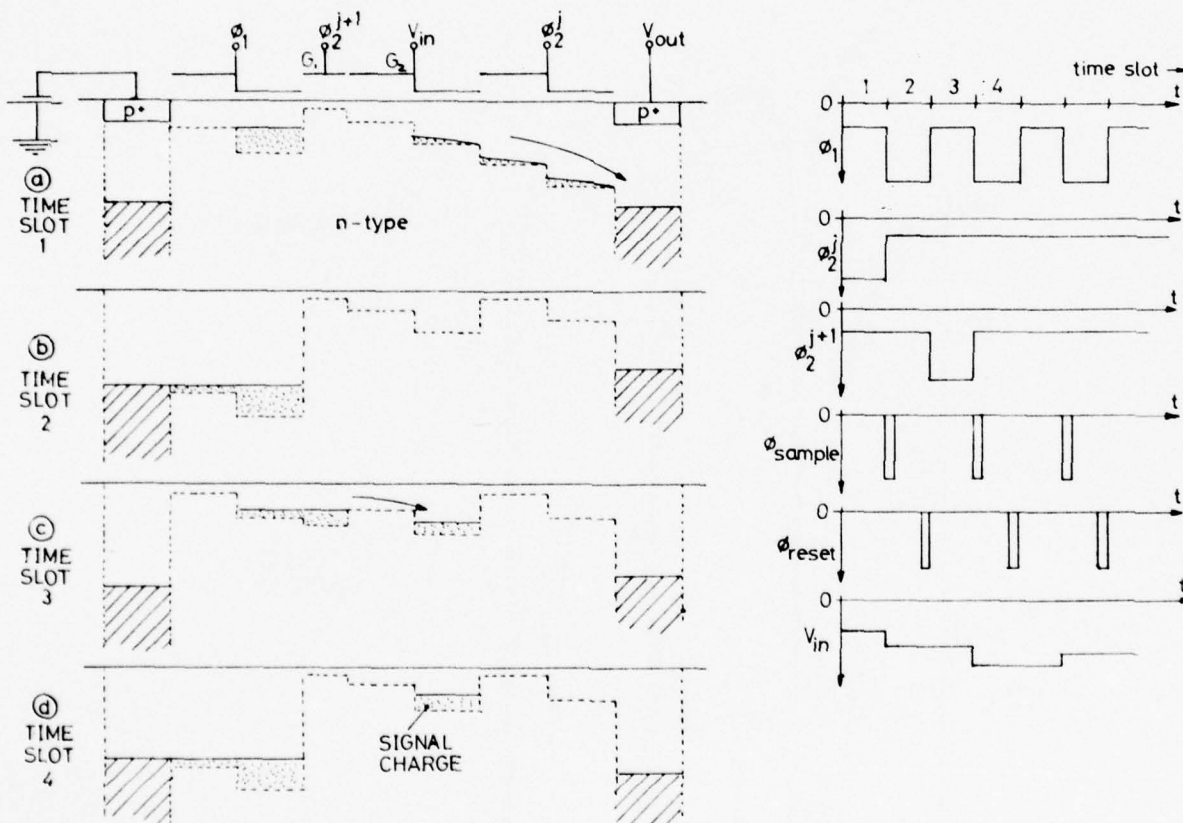
Figure 5   A cross-sectional view of each storage site showing its operation and the relevant clocking waveforms.
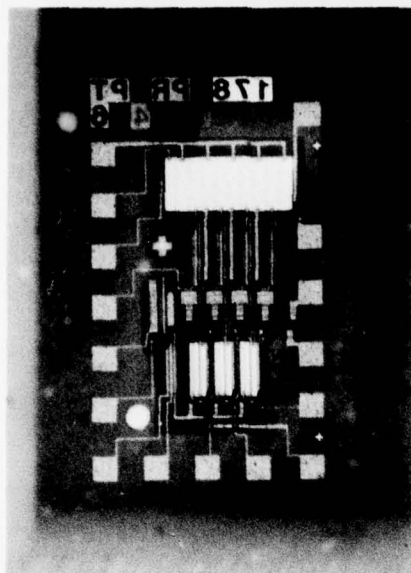
Figure 6   A photomicrograph of the prototype recursive integrator. The RGR is the feature at the top, with the inverters and storage areas below.
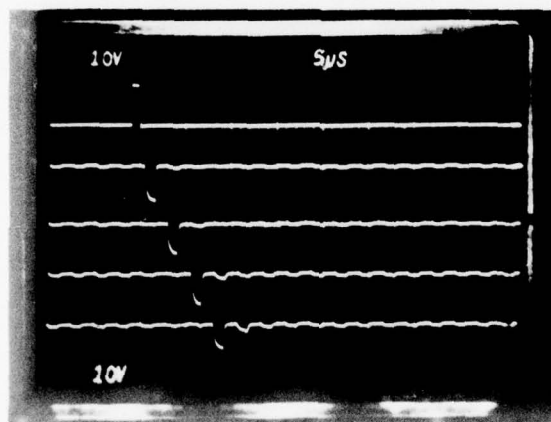
Figure 7   The impulse response of the RGR and inverters. The input pulse is shown at the top of the picture, and the inverter outputs below. Only a small charge residual can be observed. The scale is as shown.
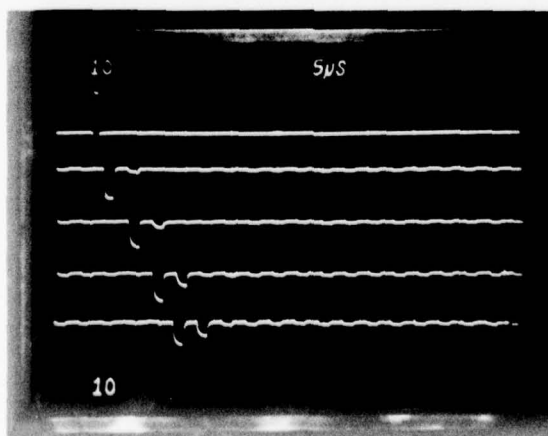
Figure 8   The bucket brigade mode of operation of the RGR with a large charge packet. The second residual is greatly enhanced, although a third does not appear. The scale is as shown.
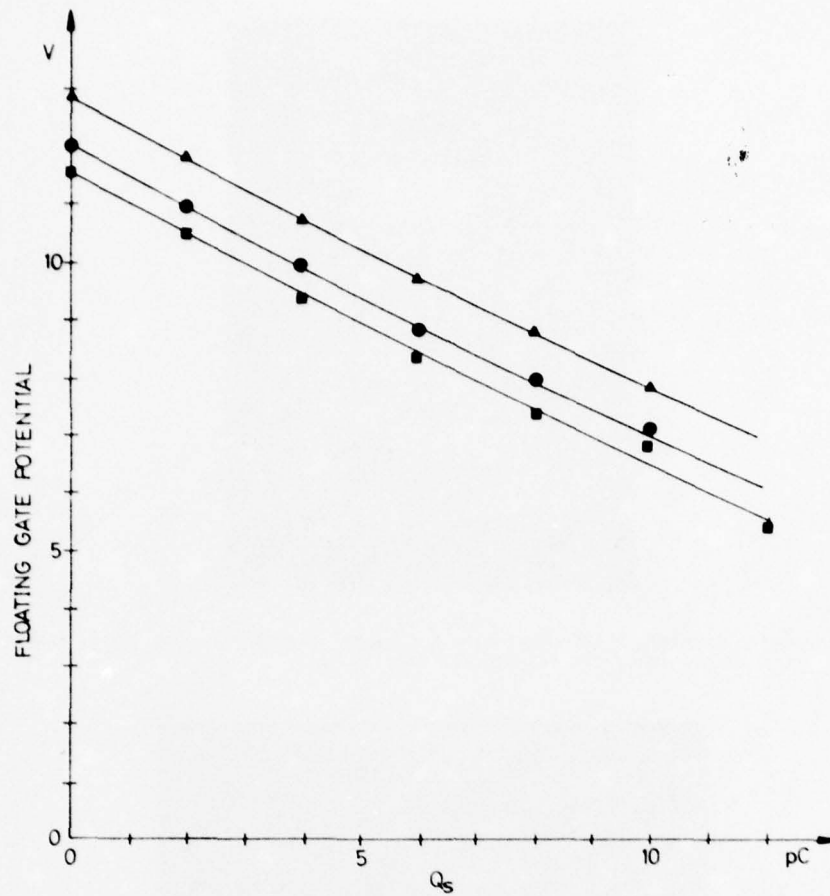
Figure 9   The transfer response of a single floating-gate of the RGR at three different bias levels.
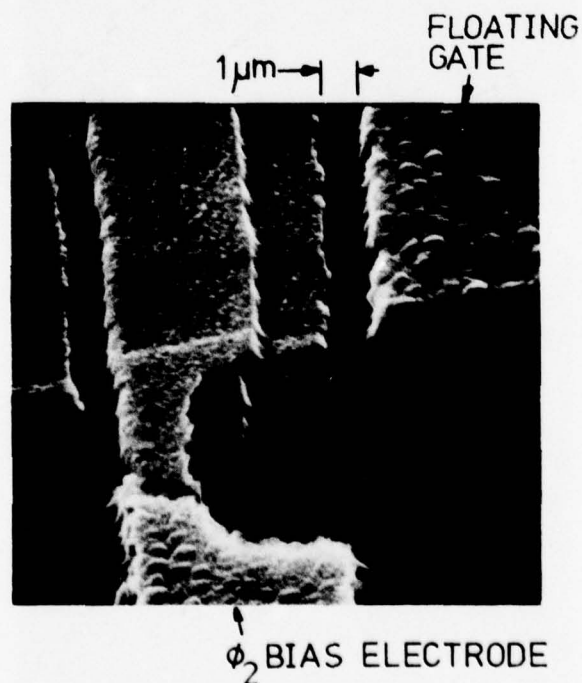
Figure 10   An electron micrograph of the split $\emptyset_2$ section of the RGR. The thick oxide bias portion can be seen separated from the thin oxide floating-gate tap by $\sim$ 1μm.
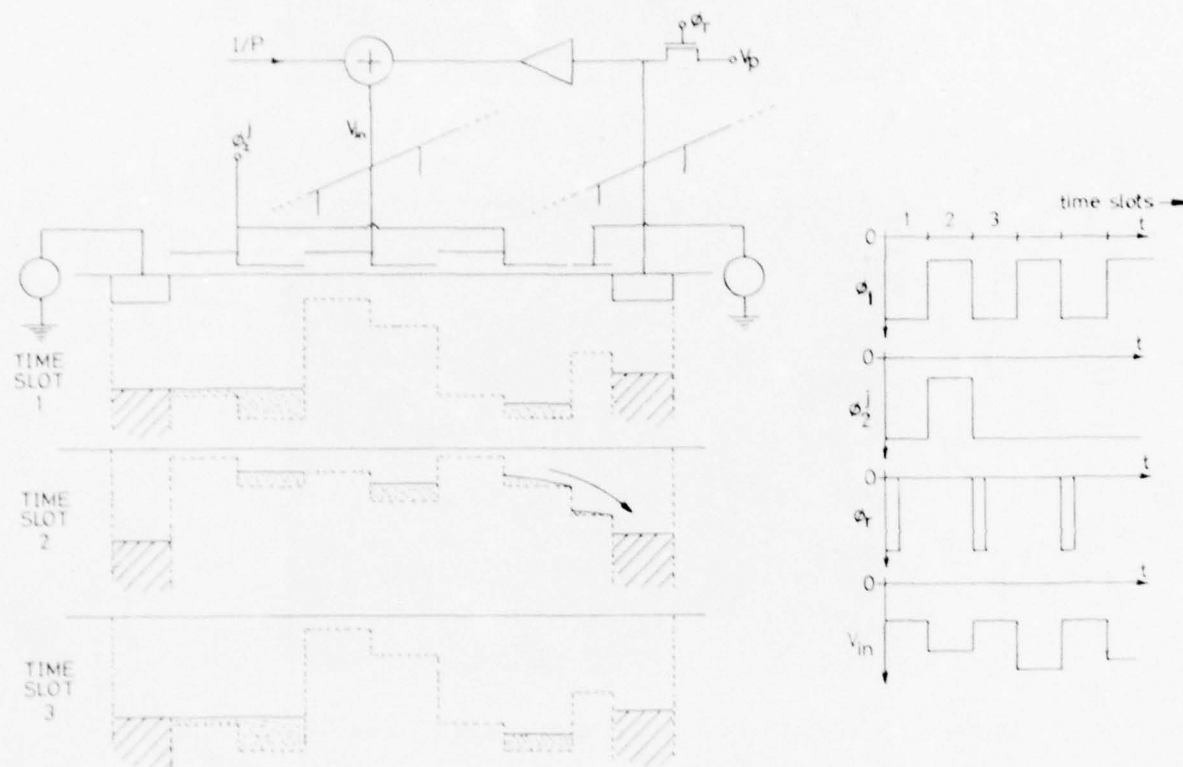
Figure 11   A cross-sectional view of the modified integrator, showing its operation and the relevant clocking waveforms. No sample/hold circuitry is required with this implementation.
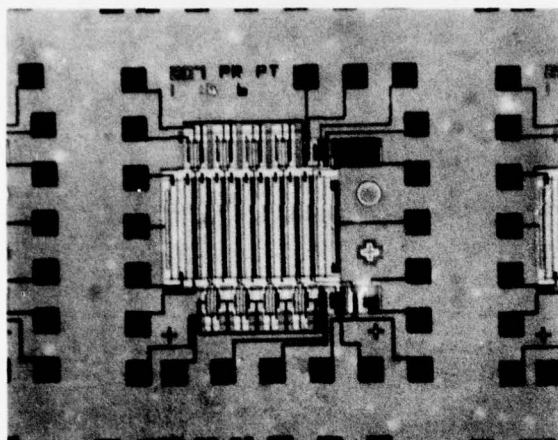
Figure 12(a)   A photomicrograph of the modified recursive integrator.
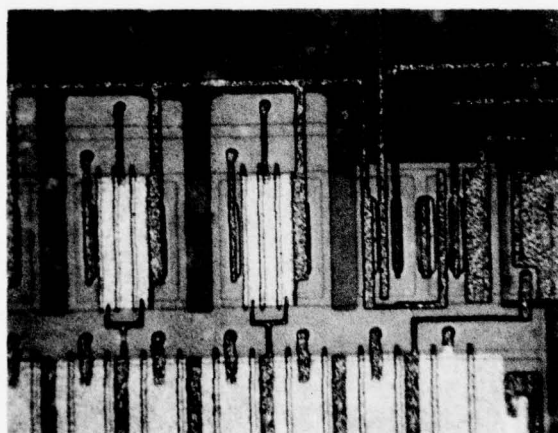
Figure 12(b)   A detailed view of the upper storage sites and output section.
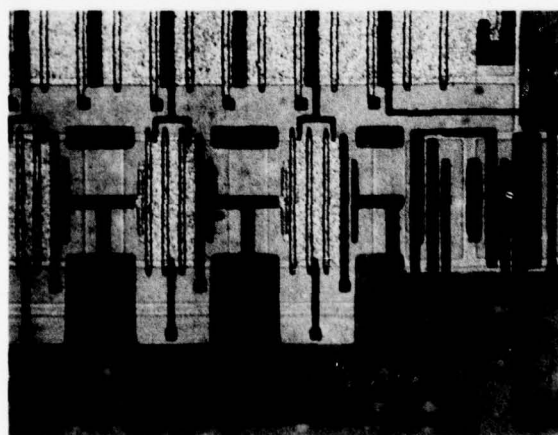
Figure 12(c)   A detailed view of the lower storage sites and output section.

LIGNE A RETARD DU TYPE CCD POUR

TRAITEMENT DU SIGNAL RADAR

APPLICATION A UN M.T.I.

J.L. BERGER
TH-CSF, D.T.E.
F38120 - ST EGREVE
FRANCE

D. RENERIC
S.T.T.A.
129, RUE DE LA CONVENTION
75015 - PARIS - FRANCE

R. VAN HOVE
TH-CSF, D.R.S.
1, RUE DES MATHURINS
92220 - BAGNEUX - FRANCE

RESUME

　　Parallèlement aux développements technologiques des CCD et notamment des lignes à retard, une étude préliminaire a été conduite sur les possibilités et limites d'un filtre MTI utilisant des CCD.

　　Pour concrétiser les résultats obtenus, le STTA a choisi la réalisation d'une maquette de filtre réjecteur de clutter utilisant un circuit CCD spécialement développé pour cette application, et devant offrir les mêmes performances que le filtre numérique.

　　La conférence décrit une ligne à retard composée de deux registres de 512 étages, sa mise en oeuvre et ses performances en efficacité de transfert, linéarité, dynamique.

　　On présente ensuite l'étude préliminaire d'un MTI et les caractéristiques du filtre désiré, les performances obtenues avec des lignes à retard (LAR) de 256 étages, les circuits associés à la LAR nécessaires pour compenser les limites du CCD, les problèmes de dérive en température.

INTRODUCTION

　　Les dispositifs à transfert de charge, commandés et contrôlés par une horloge, offrent la souplesse, la stabilité du retard, et la précision en temps du traitement numérique. Utilisant une dynamique continue en amplitude ils conservent la simplicité de fonctionnement du traitement analogique, éliminent les problèmes de conversions analogique-digitale et digitale-analogique, et offrent les avantages de taille, poids et consommation, comparés aux techniques numériques.

　　L'utilisation de ces circuits est ainsi envisagée pour de nombreuses applications dans le domaine du traitement de signal radar.

　　Afin d'évaluer les possibilités réelles de ces circuits et d'aboutir à des produits utilisables, des études de base ont été effectuées depuis plusieurs années et notamment avec le soutien de l'Armée de l'Air pour les problèmes d'entrée-sortie et les possibilités de pondération.

　　Parallèlement, une étude préliminaire sur l'emploi des CCD dans un filtrage MTI a permis d'en cerner les possibilités et les limites, de vérifier leur compétitivité réelle avec les circuits numériques et d'envisager la réalisation d'un MTI performant utilisant ces circuits. Les lignes à retard (LAR) existant en France de 128, 256, 512 étages, notamment le circuit THX 1105 de 256 étages ont permis ces premières évaluations.

　　La phase d'étude de faisabilité ayant été dépassée, l'Armée de l'Air a choisi un exemple de traitement de signal radar pour concrétiser les résultats obtenus. Ainsi le STTA fait développer un composant particulier (une ligne à retard de 2 x 512 étages en technologie canal N présentant une intégration de circuits annexes et une grande reproductibilité de caractéristiques) qui sera l'élément de base d'une maquette de filtre linéaire de réjection de clutter utilisant des circuits analogiques et devant présenter des performances identiques à celles d'un filtre numérique.

　　C'est cette ligne à retard et les résultats de l'étude préliminaire du filtre MTI qui sont présentés dans cet exposé.

1 - DESCRIPTION ET CARACTERISTIQUES D'UNE LIGNE A RETARD DE TYPE CCD POUR TRAITEMENT DE SIGNAL RADAR

1.1 - Description d'une ligne à retard CCD de 2 x 512 éléments

　　Le circuit THX1112 se compose de deux registres CCD de 512 éléments chacun, disposés sur une puce de silicium de dimensions 3,7 x 2,5 mm2.

　　Les deux registres sont identiques et indépendants. La figure 1 donne une représentation schématique du dispositif. Chacun des registres comporte :

- Un étage d'entrée où le signal est introduit sous forme de paquets de charges grâce à une conversion tension-charge. Cet étage se compose d'une diode injectrice commandée par une phase Ø1, d'une grille de contrôle G1 reliée à un potentiel référence et d'une grille G2 sur laquelle est appliquée la tension signal. Le couplage entre G1, G2 est réalisé par une zone diffusée.
  L'injection se fait selon le mode équilibrage des potentiels qui assure une bonne linéarité entre la quantité de charge introduite et la tension signal (TOMPSETT, M.F, 1976) (SEQUIN, C.H, 1975) (Mc CAUGHAN, D.V, 1976).

- La partie registre CCD proprement dite qui se compose de 4 tronçons de 128 étages chacun disposés en serpentin. Cette disposition permet d'éviter une puce de trop grande longueur. Le pas d'un étage est de 24 μ. L'avance des charges est commandée par deux signaux en opposition de phase Ø1 et Ø2.

- Un étage de sortie qui après conversion charge-tension permet d'obtenir le signal retardé sur une faible impédance.
  Cet étage de sortie se compose d'une diode collectrice dont le potentiel est périodiquement restauré par un T MOS de précharge. Le potentiel de cette diode est lu par un T MOS monté en source suiveuse réalisant une adaptation d'impédance. Le parasite d'horloge est atténué par un échantillonneur et maintien monté sur le même substrat. Le signal est disponible avant et après cet échantillonneur

et maintien.

Le circuit est réalisé dans une technologie MOS double grille silicium et canal N.

Le montage est fait dans un boîtier DUAL IN LINE 16 passages. La figure 2 donne une microphotographie du circuit.

1.2 - Mise en oeuvre

Le schéma de mise en oeuvre pratique des deux registres utilisés en mode multiplexé est donné figure 3. Deux sources d'alimentations V_{DD} et V_{SS} sont utilisées, les autres niveaux Vg1 et Vg2 sont générés à partir de ces deux sources.

Les échantillonneurs d'entrée de chaque registre, commandés l'un par Ø1 (registre A) l'autre par Ø2 (registre B) réalisent l'aiguillage des échantillons sur les deux voies parallèles. En sortie, la connexion des points SA et SB permet une bonne réjection des signaux d'horloge et réalise la reconstruction du signal qui est restitué avec une fréquence d'échantillonnage 2 Fc.

La prise d'échantillon se fait sur le flanc montant du signal d'échantillonnage et la sortie se fait sur le niveau bas de ce même signal, ce qui donne un retard exact de 512,5 Tc sur chacune des voies.

Les conditions de fonctionnement typiques sont données ci-dessous :

- V_{DD}.................... + 12 à + 16 V   - V_{SS}.................. - 5 V
- Vg1..................... + 3 V   - Vg2.................. 4,5 à 5,5 V

Vg2 est la composante continue du signal d'entrée qui doit être ajustée pour obtenir la meilleure condition d'injection : amplitude de signal maximale pour une distorsion donnée.

1.3 - Performances obtenues sur un registre simple

Les caractéristiques mesurées sur un registre pour une fréquence d'horloge de 1 MHz sont résumées ci-dessous :

$$\epsilon = \text{Inefficacité de transfert} = \frac{\text{charge perdue à chaque transfert}}{\text{charge transférée}} = 2.10^{-4}$$

$$D = \text{Dynamique} = \frac{\text{Signal maximal en sortie pour une distorsion harmonique 0,5 \%}}{\text{Signal égal au bruit mesuré avec une largeur de bande de 5 kHz}} = 60 \text{ dB}$$

A = Atténuation entrée - sortie : 9 dB.

La figure 4 illustre la distorsion observée en sortie par le relevé des niveaux d'harmonique 2 et 3 par rapport au niveau du signal, en fonction de la composante continue sur le signal d'entrée (Vg2). Dans une plage utile pour l'injection de l'ordre du volt, l'harmonique 2 se trouve à 50 dB et le niveau d'harmonique 3 se trouve 10 dB plus bas.

L'évolution des performances en fonction de la fréquence est donnée par les figures 5, 6, 7.

La figure 5 illustre l'accroissement de l'inefficacité de transfert avec la fréquence. Pour une fréquence d'horloge de 10 MHz, on observe une inefficacité de 5.10^{-4} qui devient voisine de 10^{-3} à 20 MHz. Cette inefficacité réduit l'augmentation de la bande passante lorsque la fréquence d'horloge fc augmente.

La figure 6 donne l'évolution du produit retard x bande passante avec la fréquence d'horloge. La valeur théorique de 256 est atteinte en basse fréquence (Fc = 100 kHz) ; on obtient un produit de l'ordre de 200 à 1 MHz et 100 à 10 MHz.

La figure 7 donne l'atténuation d'un signal de fréquence fc/2 (fréquence de Nyquist) par rapport à un signal BF, en fonction de fc. Cette atténuation est de 3,5 dB à 1 MHz et 9 dB à 10 MHz.

La réalisation de ce dispositif dans une technologie de transfert en volume devrait permettre d'améliorer sensiblement les performances en fonction de la fréquence.

L'évolution des performances de dynamique en fonction de la température est illustrée par la figure 8.

Chaque élément de retard est le siège d'un courant thermique Io qui double environ tous les 10 degrés et qui apporte une charge parasite Qo = IoTc, Tc étant la période de l'horloge. Après N transferts, la dérive thermique devient NQO, ce qui se traduit par une limitation de la dynamique variable avec la température.

La figure 8 donne les caractéristiques de transfert obtenues à fc = 1 MHz pour deux températures (25 et 70 C) et deux valeurs de la tension référence à l'entrée.

On relève une tension d'offset de 35 mV entre 25 et 70 C, ce qui se traduit par une réduction relative de dynamique de l'ordre de 1 dB.

Le temps de stationnement du signal dans la ligne étant de 512 μs, à cette fréquence, on en déduit un taux d'offset thermique de 70 mV/ms à 70 C.

Enfin la figure 9 donne les spectres de signal de sortie obtenus pour deux types de fonctionnement : ligne à retard seule et mode multiplexé. La fréquence d'horloge est de 5 MHz et le signal d'entrée à 1 MHz.

Le montage utilisé est celui de la figure 3. Dans le premier cas, la sortie est faite sur le point SA seul. Les niveaux d'harmonique 2 et 3 sont inférieurs à 50 dB et la bande passante maximale est de 2,5 MHz.

Dans le second cas, obtenu en reliant SA et SB, du fait d'une légère dissymétrie entre les deux registres le niveau d'harmonique 2 est à - 43 dB ; la réjection de la fréquence image autour de Fc est de - 45 dB. La bande passante maximale est alors de 5 MHz. On peut noter l'affaiblissement du signal d'horloge supérieur à 30 dB du fait de la compensation obtenue en couplant SA (parasitée par Ø1) et SB (parasitée par Ø2).

2 - ETUDE PRELIMINAIRE D'UN MTI UTILISANT CETTE LIGNE A RETARD

2.1 - Rappels

De nos jours, en aval de la tête radar, il existe deux classes distinctes de dispositifs de traitement numérique.

- La première qui effectue l'exploitation des données s'accomode parfaitement de l'organisation des calculatrices universelles.

- La deuxième qui effectue le filtrage et le traitement du signal requiert généralement des calculatrices spécialisées car ces unités travaillant en temps radar, les calculs doivent être effectués en un temps inférieur au quantum distance utilisé. Ces dispositifs sont donc destinés à des tâches nécessitant un calcul intensif caractérisé par des formules mathématiques répétitives portant sur un volume important de données.

Le filtrage anticlutter pour radar de veille qui demande un traitement continu sur un très grand nombre d'échantillons en représente une application typique.

Il y a dix ans l'avènement des registres à décalage numériques MOS a apporté une transformation complète des techniques et des équipements MTI.

Aujourd'hui tout récepteur radar de veille comporte un dispositif de réjection du clutter de sol et c'est cet équipement qui offre l'application potentielle la plus intéressante des dispositifs à transfert de charge, immédiatement disponibles et utilisables.

Le fait de conserver les signaux échantillonnés dans leur forme analogique originale, c'est-à-dire non quantifiés en amplitude, élimine l'ensemble de conversion analogique-digitale et permet de remplacer les coûteux multiplieurs numériques par de simples atténuateurs de tension résistifs.

L'utilisation des CCD représente donc une participation non négligeable dans la recherche de réalisations plus économiques, moins volumineuses et à plus faible dissipation thermique.

Pour ces raisons, de nombreuses communications ont traité des applications potentielles des CCD au domaine radar, et en particulier dans la réalisation de filtres linéaires de réjection du clutter, lorsque celui-ci exhibe un spectre Doppler de faible largeur par rapport à la fréquence de répétition du radar, et qu'il est centré au voisinage de cette fréquence et de ses harmoniques (BOUDEN. JE ..,1974) (LOBENSTEIN. H... 1975) (WARDROP. B,..1976) (KOTHMANN. W, 1976).

Toutefois, jusqu'à présent à notre connaissance, les dispositifs mis en oeuvre avec des registres à décalage analogiques, CCD ou BBD, ne comportent qu'un nombre relativement faible de cases distance dans le cas d'un filtre récursif. Inversement, s'il s'agit d'une structure multiplexe, de façon à augmenter artificiellement le nombre de quanta, le filtre est constitué essentiellement par la mise en cascade de simples annulateurs du premier ordre.

D'autre part, la dégradation des signaux et la diaphonie entre cases distance voisines, dues à une efficacité de transfert insuffisante, limitent les possibilités de traitement, en aval du M.T.I, effectué sur plusieurs cellules de résolution distance consécutives. Cet inconvénient subsiste lorsque l'on utilise des composants du type S.A.M. (Serial Analog Memory de Reticon) afin de réduire la diaphonie (KOTHMANN,W,1976).

Cet exposé a pour objet de présenter les études de définition et de conception, concernant la réalisation d'un filtre linéaire de réjection du clutter qui utilise essentiellement les dispositifs à transfert de charge THOMSON-CSF décrits précédemment.

Le dispositif envisagé correspond, sur chaque composante vidéo, au multiplexage d'au moins 1000 canaux. Chacun d'eux effectue, par exemple une fonction de filtrage à deux pôles et trois zéros, tout en maintenant une isolation d'au moins 40 décibels entre chaque canal. L'atténuation du clutter est d'au moins 50 décibels dans la mesure où cette performance est compatible avec la réjection théorique du filtre.

L'objectif principalement visé est de permettre, en sortie MTI :

- soit, d'effectuer une corrélation sur la durée du signal émis,

- soit, d'obtenir une détection à fausse alarme constante à partir d'une fonction de normalisation.

2.2 - Caractéristiques techniques du filtre MTI

Les performances globales d'une station radar équipée d'un filtre MTI linéaire dépendent non seulement du type de filtre et du "Staggering" utilisés mais également des caractéristiques des signaux parasites qui perturbent la visibilité des cibles utiles.

De façon à tester valablement ce sous-ensemble radar par comparaison instantanée, au cours d'expérimentation en vraie grandeur, nous nous proposons de réaliser en analogique, une maquette de caractéristiques et performances extrêmement voisines de celles d'un récepteur moderne pour radar de veille plane, comportant un filtre MTI numérique déjà développé depuis quelques années.

Il s'agit d'un filtre à trois mémoires, travaillant donc sur quatre impulsions, et comprenant trois zéros et deux pôles conçus pour réaliser trois types de fonctions de filtrage à savoir :

1 - Un filtre adapté aux radars de veille à relativement grand nombre N d'échos par but (N > 15) Il comporte trois zéros à l'origine (triple annulation) avec deux boucles de contre-réaction, ce qui lui confère une fonction de transfert presque rectangulaire. Il assure ainsi une bonne élimination des échos fixes si leur spectre n'est pas trop large (N > 10) avec une visibilité maximale et uniforme.

2 - Un filtre correspondant à un radar à peu d'échos dans le lobe à 3 dB (N < 15). Par rapport au filtre précédent, deux des trois zéros ont été décalés de manière à élargir la zone de réjection et présenter une meilleure atténuation autour de f = 0 lorsque le spectre des échos est plus large (N petit).

3 - Un filtre de mauvais temps présentant une bande passante étroite et une bande atténuée importante. Il s'agit dans ce cas d'un filtre non récursif à triple annulation.
Dans les radars à impulsions l'élimination des échos fixes entraîne la création de zones de réjection situées aux multiples de la fréquence de récurrence.
De façon à s'affranchir de ce problème dit des "vitesses aveugles" on effectue un changement de période d'impulsion à impulsion (Staggering). En règle générale, dès que la réjection théorique du filtre dépasse 40 dB, c'est le "staggering" qui limite les performances et dans ces conditions on adopte la forme rectangulaire des filtres 1 et 2. Si au contraire les paramètres sont tels que la fréquence centrale du spectre de clutter représente une partie notable de la fréquence de récurrence c'est le filtre qui limite les performances en réjection et on peut dans ce cas se contenter du filtre 3 non récursif à 3 mémoires.

Les trois éléments de retard sont associés de façon à former deux filtres élémentaires en cascade (décomposition en éléments simples).

- un filtre simple annulation
- un filtre à deux retards à entrées pondérées et comportant deux boucles de contre-réaction.

Le synoptique du filtre MTI est représenté figure 10.
La transformée en Z donne immédiatement :

$$H(Z) = (1 - Z^{-1}) \frac{(1 - 2AZ^{-1} + Z^{-2})}{(1 - R'Z^{-1} + RZ^{-2})}$$

et les fonctions de transfert des trois filtres sont représentées (figure 10).

Cet équipement est destiné à être associé à un radar de veille plane répondant aux caractéristiques suivantes :

- Largeur de l'impulsion émise .... Γ = 1,75 microseconde
- Fréquence de récurrence .... Fr = 750 Hz (± 10 % de STAGGER)
- Portée MTI .... de 0 à 80 Nautiques
- Nombre d'échos dans le lobe d'antenne .... N = 20 environ

Après amplification avec une bande passante adaptée à la largeur Γ = 1,7 μs de l'impulsion émise (soit environ 600 kHz), les signaux à fréquence intermédiaire issus du préamplificateur radar sont transposés autour d'une porteuse à fréquence nulle et séparés en deux voies, en phase I et en quadrature Q. Ces signaux I et Q contenant les informations d'amplitude et de phase sont représentés en bande de base (300 kHz environ) par deux signaux vidéo bipolaires, projections sur deux axes perpendiculaires des informations d'entrée.

Ces signaux vidéo I et Q sont échantillonnés au moins une fois par largeur d'impulsion émise (1,7 μs). Un ensemble de trois retards, chacun égal à la période de récurrence (1330 μs environ) permet de disposer des informations de quatre récurrences successives pour effectuer à chaque quantum distance, le filtrage M.T.I. en peigne désiré.

En adoptant dans le cas présent une fréquence d'échantillonnage de un mégahertz, c'est-à-dire un quantum distance de traitement de une microseconde, la portée M.T.I. de 80 nautiques nécessite de traiter simultanément et continuellement environ 2000 canaux de données (1000 sur la voie I et 1000 sur la voie Q) échantillonnées à une cadence de 750 Hz, égale à la fréquence de répétition du radar.

Une ligne à retard (LAR) THX 1112, de 2 x 512 étages de transfert permet en principe de multiplexer 1000 canaux sur une voie I ou Q. En fait, comme indiqué au § suivant, de façon à augmenter artificiellement l'efficacité de transfert intrinsèque du CCD, nous utilisons deux étages de transfert de la LAR pour chaque échantillon de donnée à l'entrée. Ceci nécessite donc deux LAR, THX 1112, pour multiplexer les 1000 canaux d'une voie I ou Q. La figure 11 représente le synoptique du sous-ensemble radar et la figure 12 la version analogique du filtre M.T.I. trois mémoires.

A la sortie du filtre MTI, un traitement C.F.A.R. est utilisé pour ramener les signaux parasites, subsistant après filtrage, et s'étendant sur un certain nombre de quanta distance, à un niveau aussi constant que possible et égal à celui du bruit thermique du récepteur.

Le système du type logarithmique, différentiation, puis expansion, conçu pour normaliser la puissance du bruit ou du clutter réalise un excellent compromis entre la simplicité de réalisation et l'efficacité de la régulation (SKOLNIK I.M, 1970). Ce système revient en première approximation à diviser le signal détecté dans chaque case distance, par une estimation en moyenne géométrique, de l'ambiance entourant ce signal sur un intervalle de mesure centré sur la case intéressée.

2.3 - Performances obtenues en utilisant une LAR, THX 1105.

Afin d'évaluer les possibilités réelles et les limites des dispositifs lignes à retard à CCD dans les applications M.T.I., nous avons mené une étude préliminaire basée uniquement sur l'utilisation de LAR existant en France, notamment le circuit THX 1105 de 256 étages commandés par deux phases et comportant un étage d'entrée à deux grilles et un étage de sortie avec amplificateur tampon intégré.

Echantillonnage et maintien

Dans un ensemble de conversion analogique-digitale pour traitement numérique de signaux vidéo radar, de façon à résoudre le problème du taux de variation très élevé de ces signaux, le codeur proprement dit est précédé d'un échantillonneur-bloqueur. L'objet de ce dernier est de permettre *l'utilisation de circuits précis de codage* ayant des vitesses de réponse rapides par rapport à la cadence de codage et non par rapport aux variations des signaux à coder.

De même pour les LAR à CCD, un échantillonneur-bloqueur très performant est utilisé en entrée afin d'éviter les variations du signal durant une période d'horloge de commande des phases. La fonction d'échantillonnage propre au CCD sera réservée pour réaliser les fonctions de démultiplexage puisque le nombre de 512 éléments par registre est encore insuffisant pour couvrir la portée MTI du radar.

De plus, à cause de ce démultiplexage, la largeur de bande des signaux vidéo sera éventuellement supérieure à la fréquence d'horloge. Dans ces conditions, puisque le signal utile apparait en sortie des registres comme la modulation d'un signal parasite d'amplitude importante, il sera absolument nécessaire d'échantillonner la sortie des registres en synchronisme avec la commande des phases pour supprimer le plus efficacement possible tous les signaux indésirables de lecture des charges.

Inefficacité de transfert

La valeur non nulle de l'inefficacité de transfert ε a pour conséquence une dispersion en temps du signal vidéo, créant de ce fait, de la diaphonie entre les différents canaux multiplexés qui réalisent sur chaque composante I ou Q la fonction du filtrage M.T.I.

Si on désire en aval de ce filtrage, soit effectuer une corrélation sur la durée de l'impulsion émise, soit obtenir une détection T.F.A.C. à partir d'une fonction de normalisation, il faut maintenir au minimum, une isolation de 40 décibels entre chaque canal.

Sur une impulsion isolée d'amplitude unité, la dispersion peut être chiffrée en première approximation, par le théorème du développement du binôme :

$$\left[(1 - \varepsilon) + \varepsilon\right]^N = \sum_{p=0}^{N} \frac{N!}{p!(N-p)!} (1 - \varepsilon)^{N-p} (\varepsilon)^p$$

Après N transferts, ceci se traduit par :

- une atténuation A = exp (-Nε) pour l'impulsion

- la création de plusieurs impulsions résiduelles dont la première a pour amplitude Nε , la seconde $\frac{(N\varepsilon)^2}{2}$, la troisième $\frac{(N\varepsilon)^3}{6}$ ... etc.

Dans un registre à deux phases de n = 512 étages, le nombre de transferts N = 2 n $\approx$ 1024. Les CCD, THX1112 possèdent des ε chiffrés entre 2 10^{-4} et 5 10^{-4} pour des fréquences de 1 à 5 MHz, soit Nε compris entre 0,2 et 0,5.

Ainsi en supposant un réglage de gain réalisant l'annulation parfaite, d'impulsions vidéo simulant un écho fixe, entre la voie directe et la voie retardée comme par exemple dans un filtre simple annulation, la réjection ne dépasserait guère 10 dB puisqu'il resterait les impulsions résiduelles dont la première serait encore de 20 à 30 % de l'impulsion signal.

Il s'agit donc de gagner artificiellement, à l'aide de circuits associés, un facteur de 20 à 50 sur l'efficacité de transfert intrinsèque des CCD.

La figure 13 décrit une telle *méthode d'amélioration de l'efficacité de transfert équivalente*, utilisant d'**une part**, un étage de transfert redondant sur deux, de façon à réduire au second ordre $\frac{(N\varepsilon)^2}{2}$ les impulsions résiduelles, d'autre part une annulation supplémentaire réalisée à l'aide d'un étage de transfert annexe et d'un amplificateur différentiel qui abaisse théoriquement au 4ème ordre l'effet résiduel.

Les photos de la figure 14 illustrent les résultats obtenus à l'aide d'un circuit THX 1105 à 1 MHz de fréquence de commande de phases, sur signaux rectangulaires. Le mode d'introduction du signal est dans ce cas la méthode d'équilibrage des potentiels qui provoque une atténuation entrée-sortie de 20 dB environ. Les différentes tensions d'alimentation ne sont pas particulièrement optimisées mais réglées de façon à illustrer le plus lisiblement possible l'amélioration apportée aux différentes étapes du procédé.

La photo de la figure 14-f illustre de manière plus exacte la sortie directe du TX 1105 pour une fréquence légèrement supérieure à 1 MHz et un mode d'injection à l'entrée différent (l'impulsion d'échantillonnage est appliquée sur la première grille d'entrée au lieu d'être appliquée sur la diode injectrice).

Cependant des mesures effectuées à l'aide d'un oscillo. et d'un tiroir différentiel (par exemple le TEKTRONIX 7704 A muni d'un tiroir 7A13) font apparaître les limites de la correction.

Si théoriquement le rapport entre le signal utile et l'impulsion résiduelle d'ordre 4 est de 60 dB, pratiquement nous n'obtenons que 50 dB, car les phénomènes mis en jeu sont sans aucun doute plus complexes que le modèle théorique choisi (développement binomial).

Linéarité

La linéarité de la tension de sortie des CCD en fonction de celle d'entrée est un paramètre fondamental car il détermine le taux d'annulation que l'on peut atteindre dans un dispositif de filtrage linéaire.

Afin d'estimer directement les performances possibles en linéarité et dynamique nous avons réalisé des filtres simple annulation avec correction d'inefficacité.

La cellule élémentaire est représentée figure 15 . La fréquence Fc de fonctionnement du CCD et le nombre d'étages du registre déterminent le temps de retard égal à $\frac{256}{Fc}$ pour le THX 1105, et l'intervalle de fréquence entre deux zéros du filtre $\Delta F = \frac{Fc}{256}$. Les résultats qui suivent ont été obtenus pour une fréquence Fc de 1,237 MHz, c'est-à-dire un retard entrée sortie CCD de 206,88 μs et des annulations de fréquence tous les 4,833 kHz.

Nous avons utilisé pour les mesures un générateur sinusoidal de grande pureté spectrale dont les harmoniques de rang 2 et 3 sont de l'ordre de - 80 dB sous le fondamental (voir photo figure 16) et nous avons vérifié, dans un premier temps que les circuits associés au CCD n'introduisaient pas d'erreurs de non linéarité.

La photo de la figure 17 présente la sortie du CCD après correction d'ε , soit -55 dB pour H_2 et environ -60 dB pour H_3.

La photo figure 18 présente le résultat en sortie filtre pour une fréquence située dans la bande passante soit -60 dB au moins pour H_2 et H_3. Enfin la photo de la figure 19 illustre l'atténuation obtenue pour une fréquence multiple de 4,833 kHz soit une annulation du fondamental supérieure à 65 dB, les raies harmoniques se maintiennent au niveau du cas précédent (-60 dB) puisque l'annulation ne peut s'effectuer que sur la raie fondamentale, seule présente à la fois dans la voie directe et dans la voie retardée.

Nous avons obtenu des résultats identiques en cascadant jusqu'à l'ordre 3 ce filtre simple annulation.

Dynamique

La dynamique est définie par le rapport entre l'amplitude maximale en sortie et l'amplitude du signal égal au bruit en sortie. L'évaluation de cette dynamique peut s'effectuer aisément à l'aide d'un analyseur de spectre de largeur de bande d'analyse très petite par rapport à la fréquence de commutation des phases de commande, de façon à ne pas prendre en compte dans la mesure les résidus d'horloge.

Nous avons effectué aussi différentes mesures du niveau de bruit à l'aide d'un voltmètre sélectif (HP 312B) dans des bandes d'analyse de 200 Hz et de 3100 Hz.

Bande	200 Hz	3,1 kHz
Appareil seul	0,8 μV	2,8 μV
CCD	3,75 μV	14,5 μV

Ces résultats vérifiant que les puissances de bruit sont dans le rapport des bandes de mesure, la tension efficace de bruit dans une bande utile de 300 kHz serait de l'ordre de $\left(14,5 \sqrt{\frac{300}{3,1}}\right)$ μV soit environ 0,14 mV.

Dans les mêmes conditions de mesure nous avons obtenu en sortie une tension signal de 40 à 50 mV pour une atténuation d'environ 60 dB des harmoniques 2 et 3, soit une dynamique possible d'au moins 50 dB.

2.4 - Dérives en température

Des mesures quantitatives (voir figure 20) ont confirmé l'influence de la température sur les variations de perte d'insertion (atténuation entrée-sortie).

L'amplitude de ces variations est incompatible avec le taux d'annulation cherché : il est donc nécessaire de les réduire.

Nous avons réalisé un système de régulation automatique de gain (figure 21) dans lequel les variations de perte d'insertion sont évaluées, en zone de traitement non utile (fin de portée radar), par injection d'un signal calibré.

Ce système a donné entièrement satisfaction ; il a permis de maintenir le taux de réjection à une valeur proche de 60 dB avec des variations de perte d'insertion pouvant aller jusqu'à 20 dB. (Cas extrême permettant de vérifier l'efficacité de régulation).

2.5 - Conclusion

L'utilisation des circuits THX 1105 immédiatement disponibles nous a permis d'aborder la phase de faisabilité d'un MTI performant et de relier le plus vite possible les paramètres spécifiques des CCD à ceux des dispositifs de traitement du signal radar.

La réalisation et les essais du sous-ensemble radar utilisant les THX 1112 permettront de déterminer quels progrès seront requis à la fois du point de vue technologique et du point de vue conception afin de rendre les CCD encore plus utilisables dans le traitement analogique du signal radar : aspect reproductibilité des performances, maîtrise des paramètres, simplicité de mise en oeuvre, intégration de circuits supplémentaires, ...etc.

Dès à présent le traitement analogique à base de CCD semble réaliser un excellent compromis en termes de coût, encombrement, poids et dissipation thermique, et en admettant une dynamique accrue de fonctionnement, 70 à 80 dB par exemple, les CCD résoudront de façon élégante et peu onéreuse le problème de conserver au rapport signal à bruit une valeur constante indépendante des signaux à traiter.

R E F E R E N C E S

1. BOUDEN JE, TUMLINSON M.J., 1974, "CCD Chebyshev filter for radar MTI applications", Electronics Letters 4th April 1974, vol.10 No 7, p 89.

2. D.V. Mc CAUGHAN, J.G. Harp, 1976, "Phase referred input : A simple new linear CCD input method", Electronics letters 9 th december 1976 vol.12, no 25, p 682.

3. W. KOTHMANN, 1976, "MTI Filters using analogue memories" AGARD-CP-1976 on new devices, techniques and systems in radar p 10-1.

4. H LOBENSTEIN, D.N. LUDINGTON, 1975 "A charge transfer device MTI implementation", IEEE 1975 International Radar Conference, p.107.

5. C.H. Sequin and A.M. Mohsen, "Linearity of electrical charge injection into charge - coupled devices", IEEE solid state circuits, vol. Sc 10, Apr. 1975.

6. M.I. SKOLNIK, 1970 "Radar handbook", Mc Graw-Hill, p 5-29.

7. M.F TOMPSETT 1976 "Using charge coupled devices for analog delay" Proc. CCD Appl. Conf.Sept 1976 - p147-150.

8. B. WARDROP, E.BULL, 1976. "The Application of charge coupled devices to moving target indicator filters", the marconi review, fourth quarter, 1976, p 201-223.

Fig.1 Organisation du registre à décalage THX 1112 de 2 x 512 étages.

Fig.2 Microphotographie du circuit THX 1112.
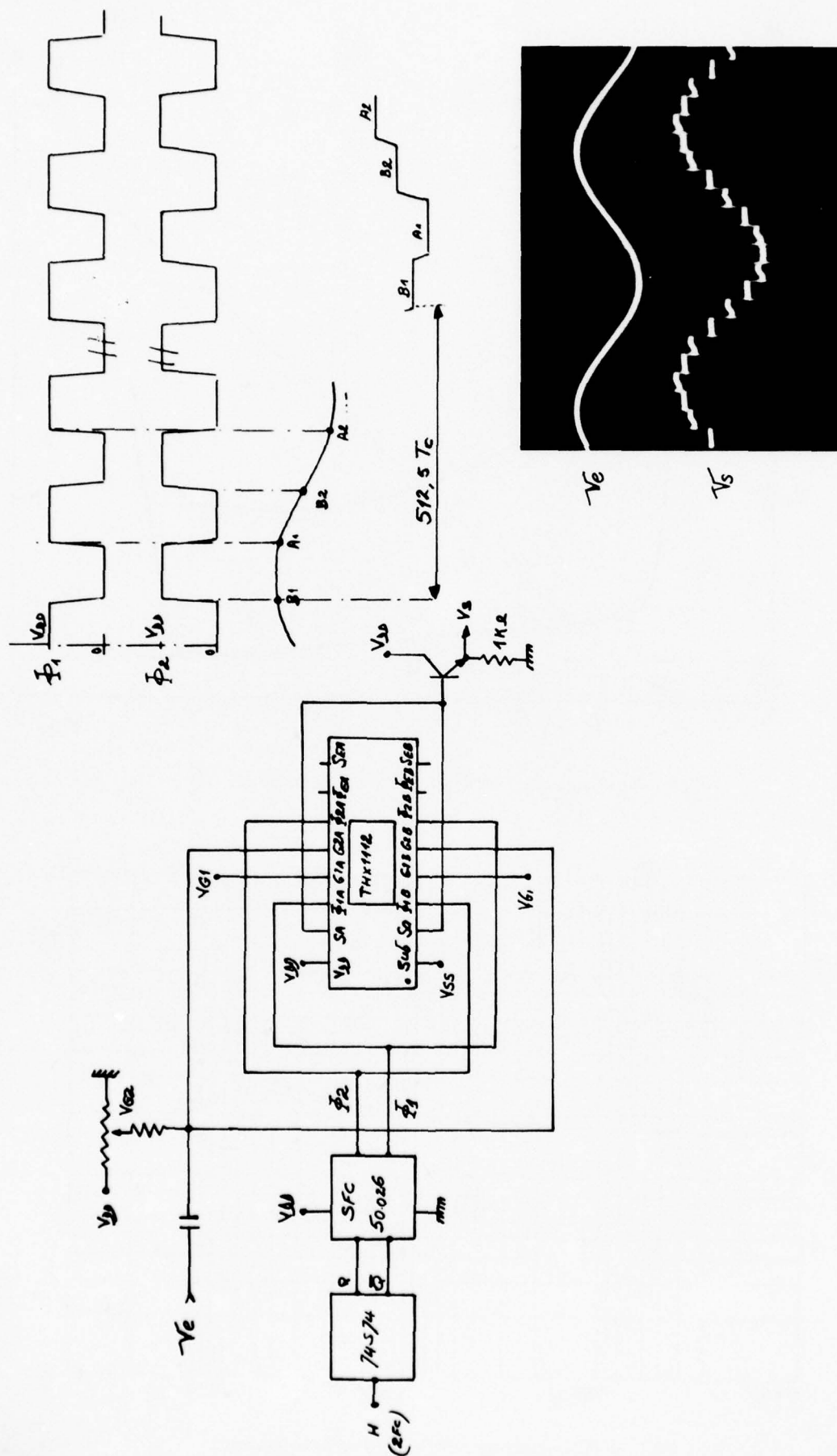
Fig.3 Schéma de mise en oeuvre en mode multiplexé.
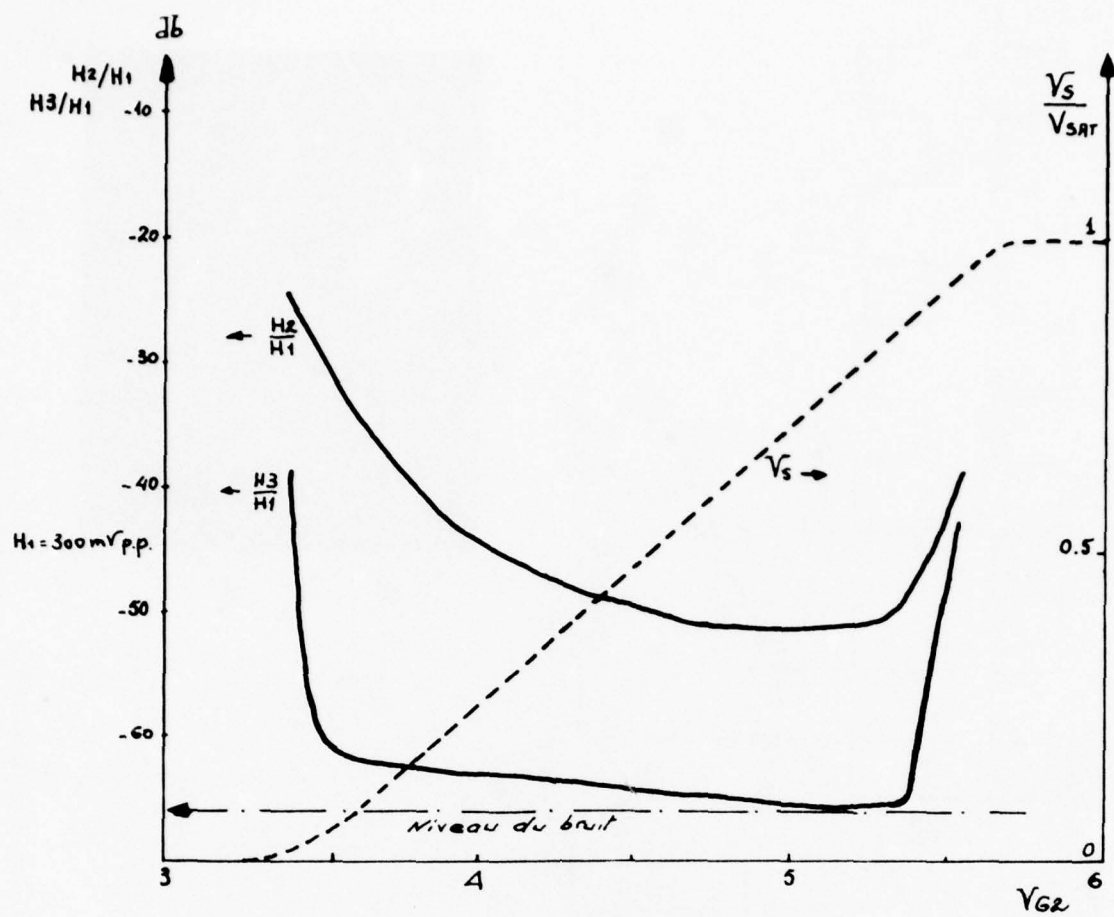
Fig.4   Distorsion: relevé harmonique 2 et 3 en fonction de VG₂
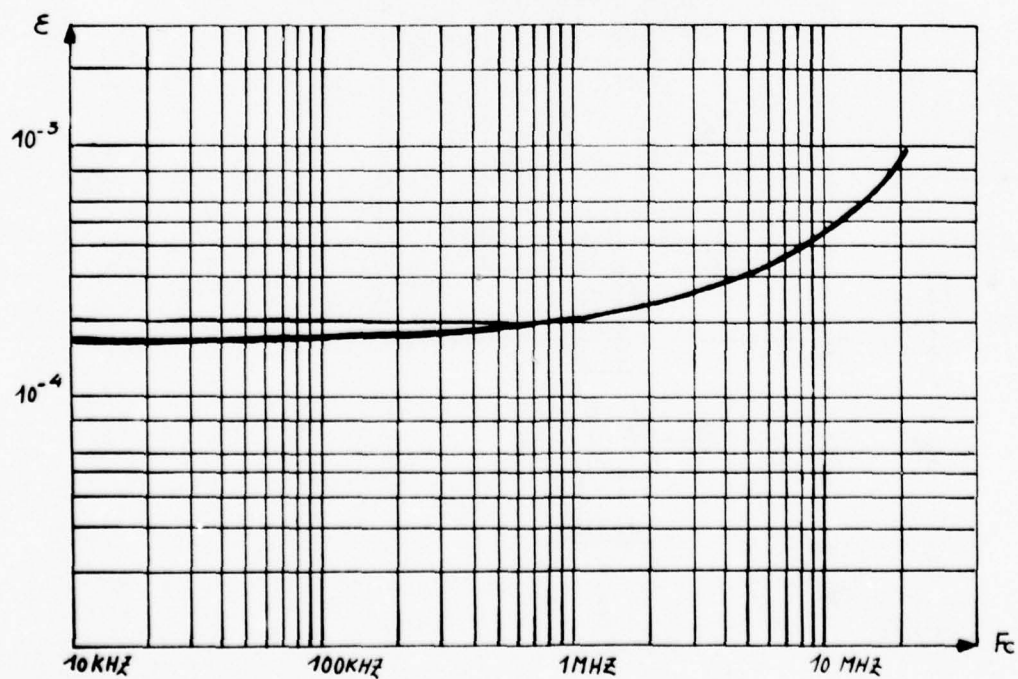
Fig.5   Inefficacité de transfert en fonction de la fréquence.
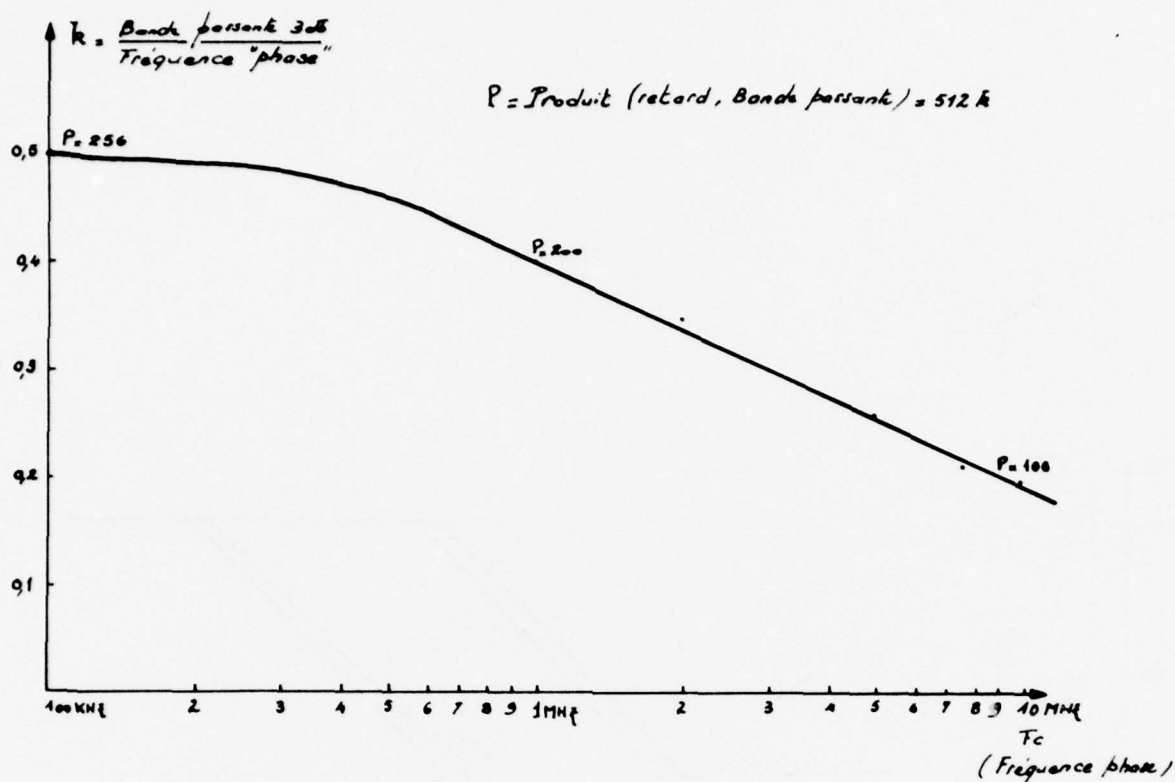
Fig.6 Evolution du produit retard x bande passante en fonction de la fréquence phase (Horloge).
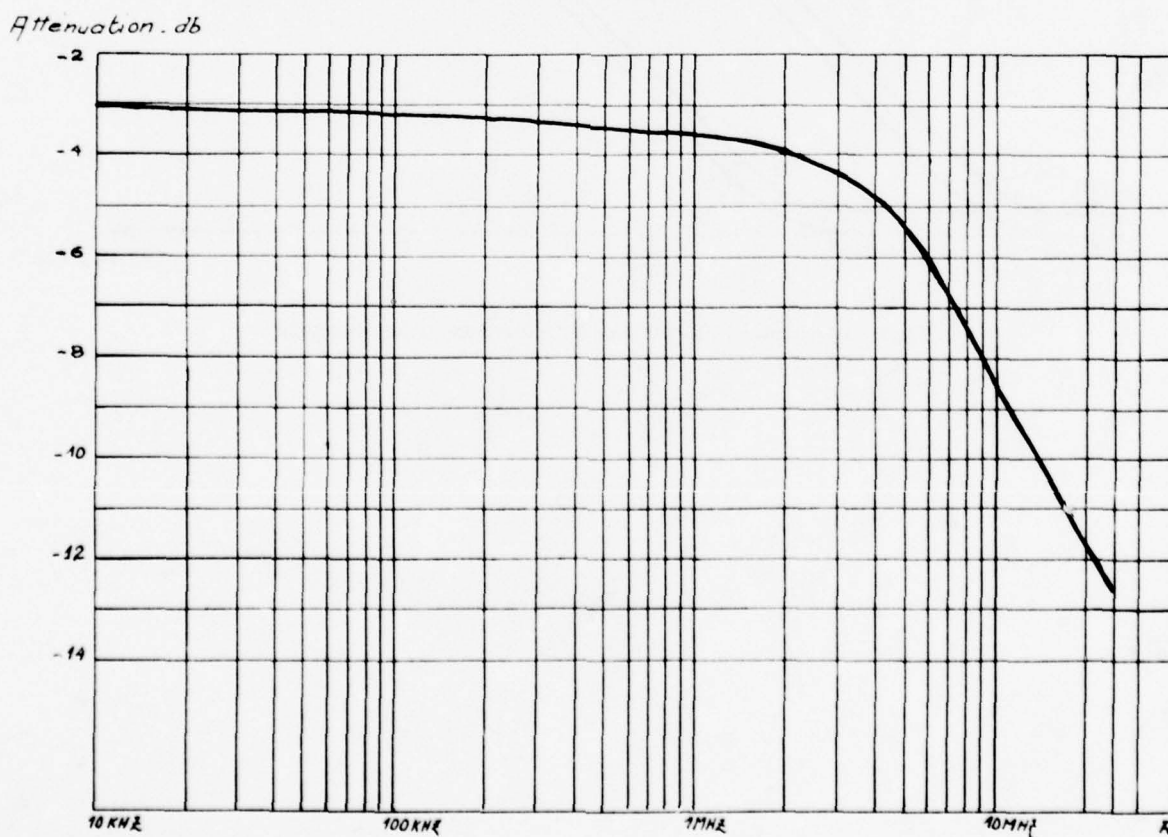
Fig.7 Atténuation à $\frac{f_c}{2}$ en fonction de f_c.
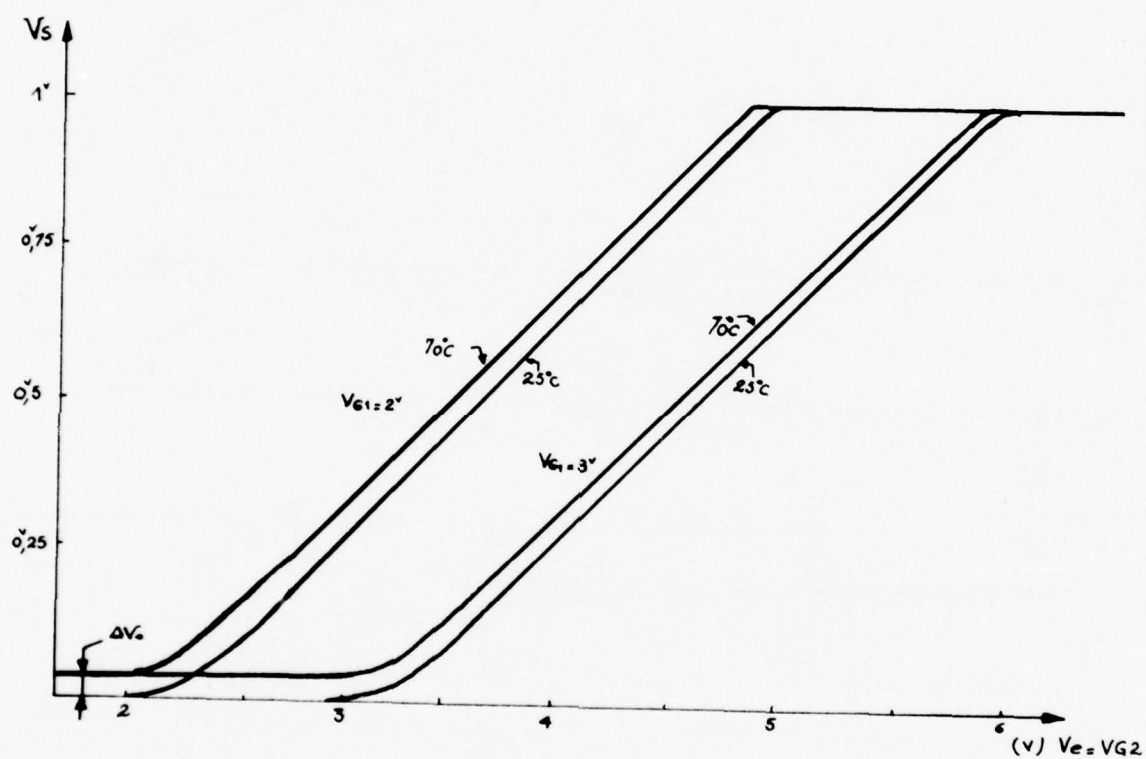
Fig.8   Dynamique en fonction de la température -offset thermique-
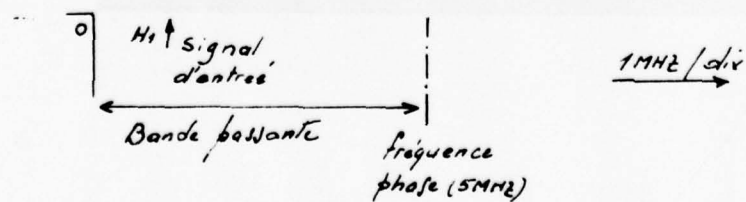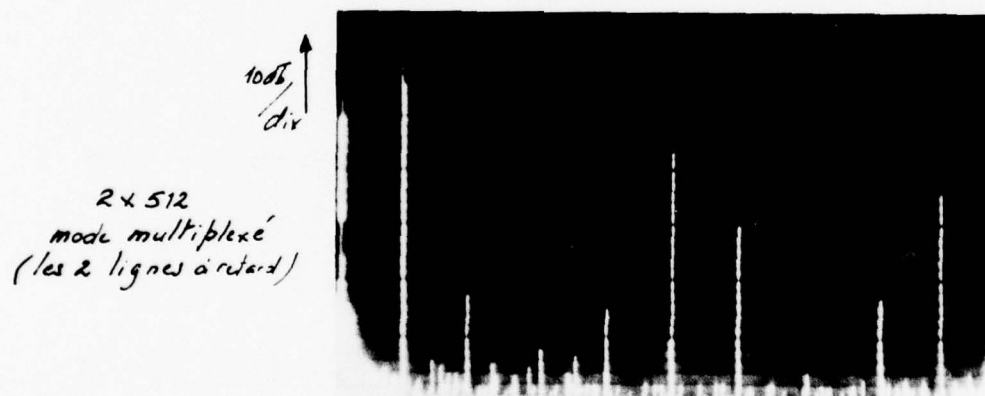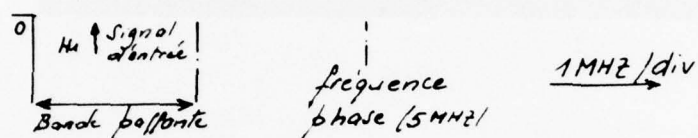
Fig.9   Spectres de sortie — 1 seule ligne à retard (1 x 512)
2 lignes à retard (2 x 512) en mode multiplexé
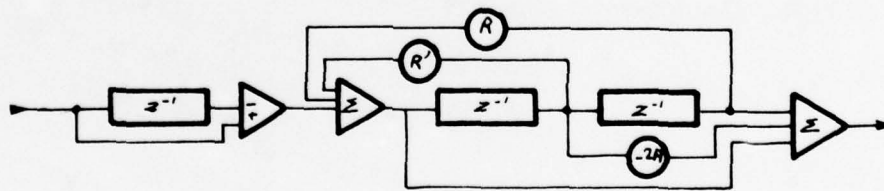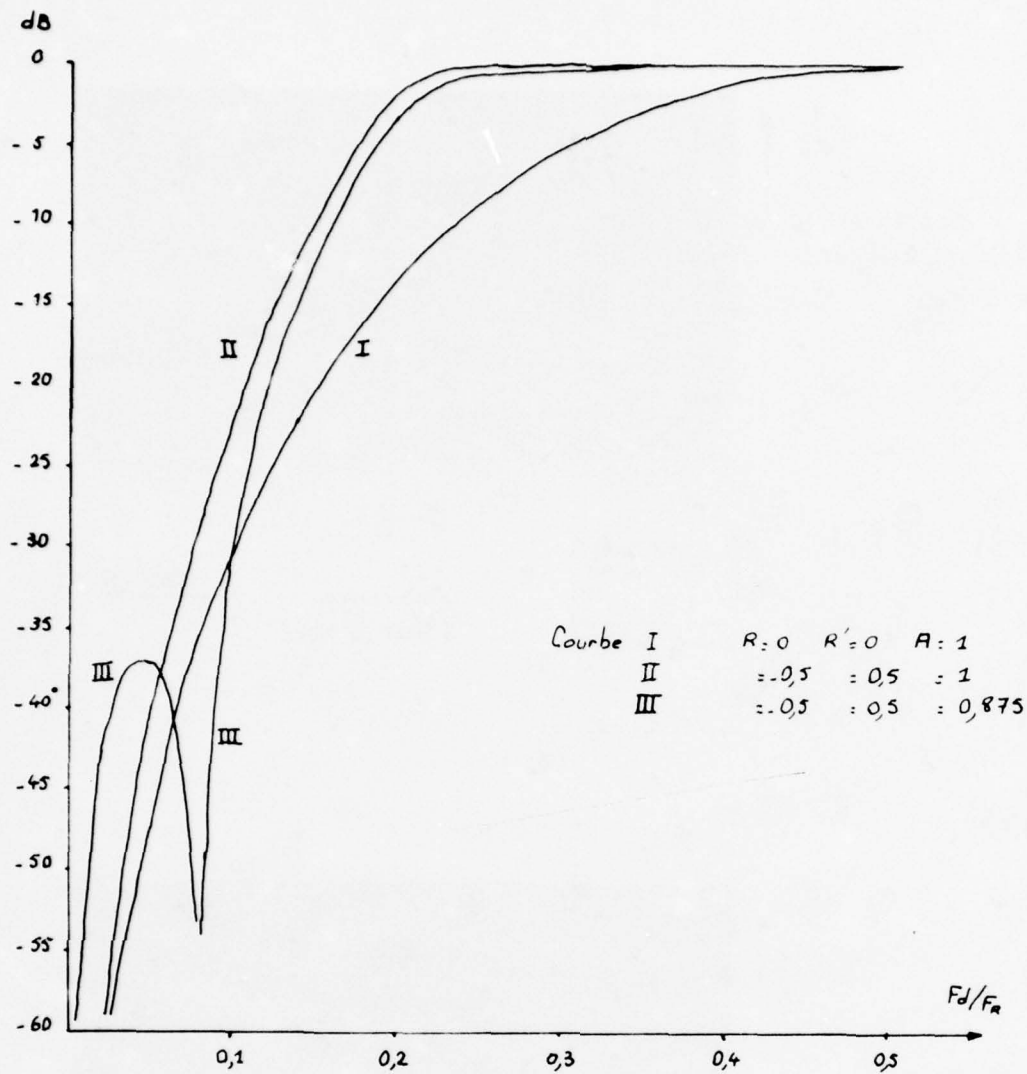
Fig.10 Fonctions de transfert, filtre 3 mémoires.
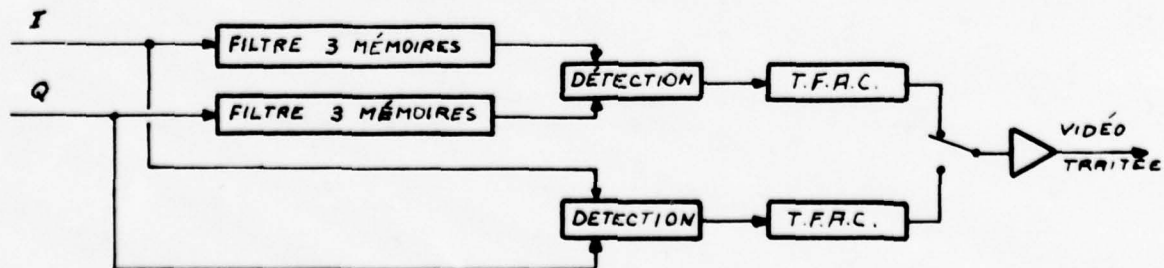
Fig.11 Synoptique de traitement.
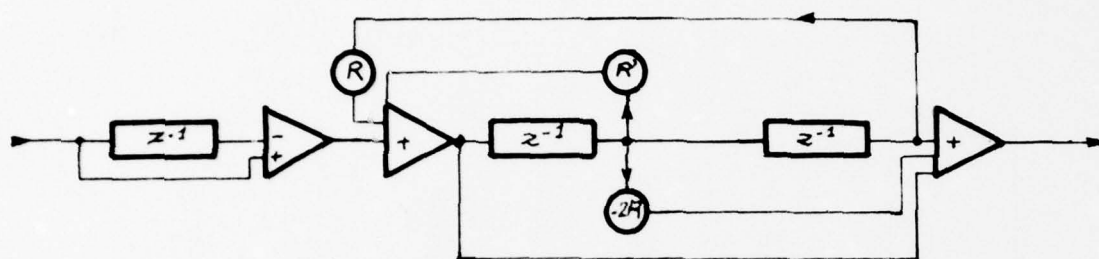
(a)
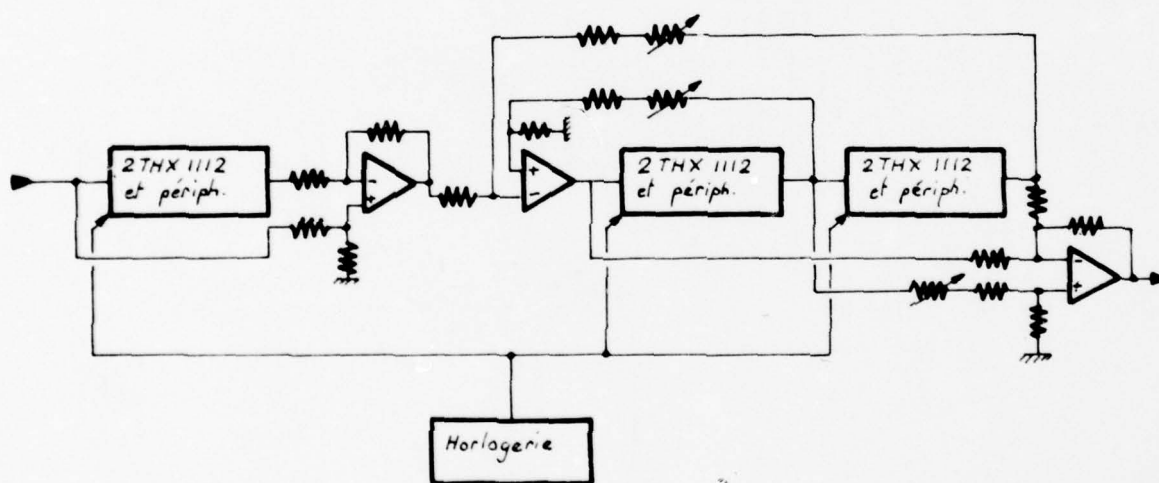
(b)

Fig.12   12a — Filtre 3 mémoires
        12b — Filtre 3 mémoires à CCD (1 composante)

Entrée f_e E.B.1 (A) $2f_e$ C.C.D. (B) f_e E.B.2 (C) f_e E.B.3 (D) (k) Σ Sortie (E) f_e E.B.1 (A')

(A)

(B) 1 $1+n\varepsilon$ $n\varepsilon + \dfrac{n^2\varepsilon^2}{2}$ $\dfrac{n^2\varepsilon^2}{2} + \dfrac{n^3\varepsilon^3}{6}$ $\simeq \dfrac{n^3\varepsilon^3}{6}$

(C) $1+n\varepsilon$ $\dfrac{n^2\varepsilon^2}{2} + \dfrac{n^3\varepsilon^3}{6}$

(D) $1+n\varepsilon$ $\dfrac{n^2\varepsilon^2}{2} + \dfrac{n^3\varepsilon^3}{6}$

(E) $1+n\varepsilon$ 0 $\simeq -\dfrac{n^4\varepsilon^4}{4}$

avec $k \simeq \dfrac{n^2\varepsilon^2}{2}$
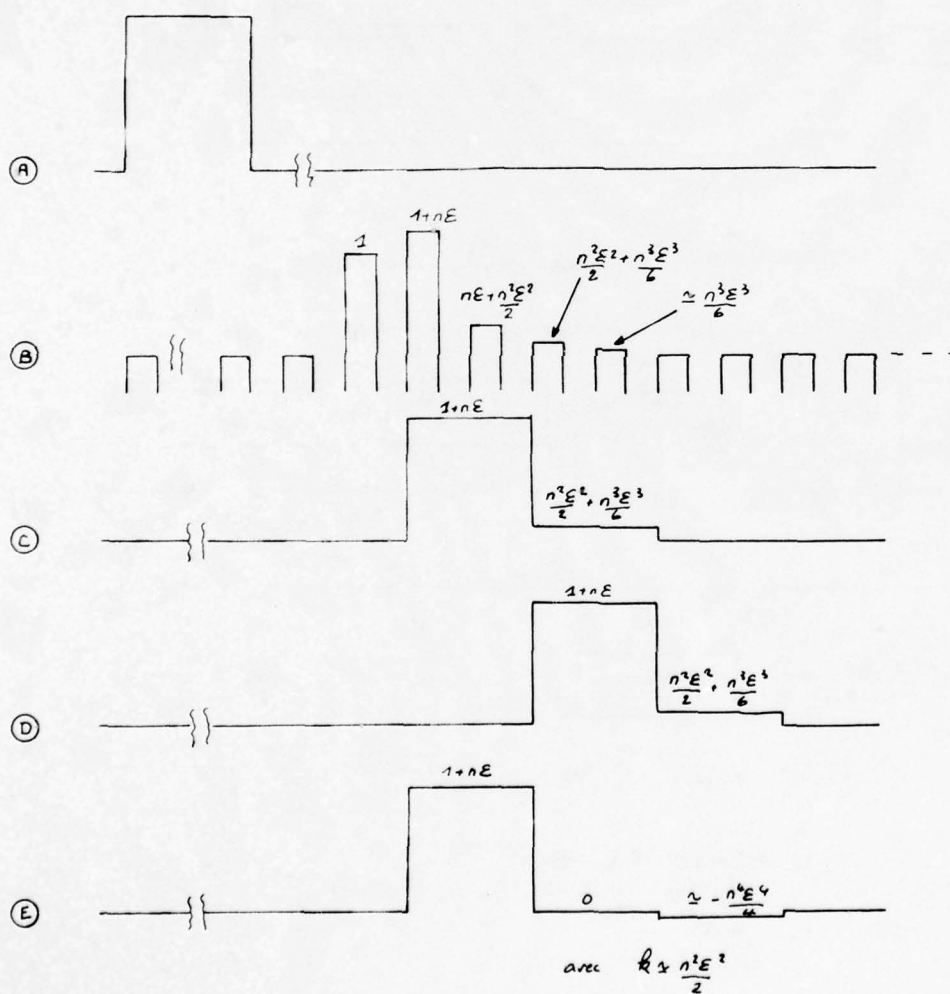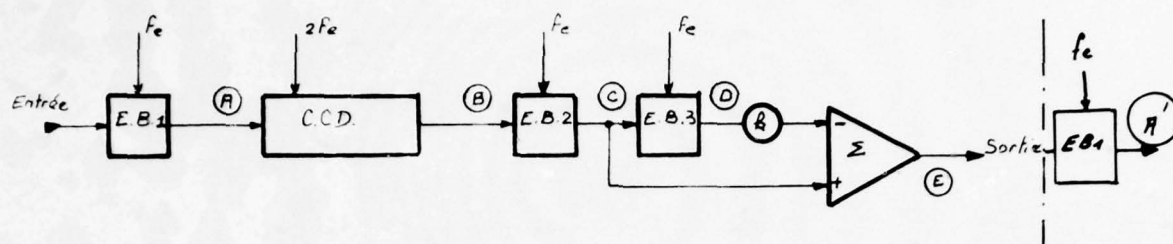
Fig.13 Amélioration de l'efficacité de transfert équivalente.
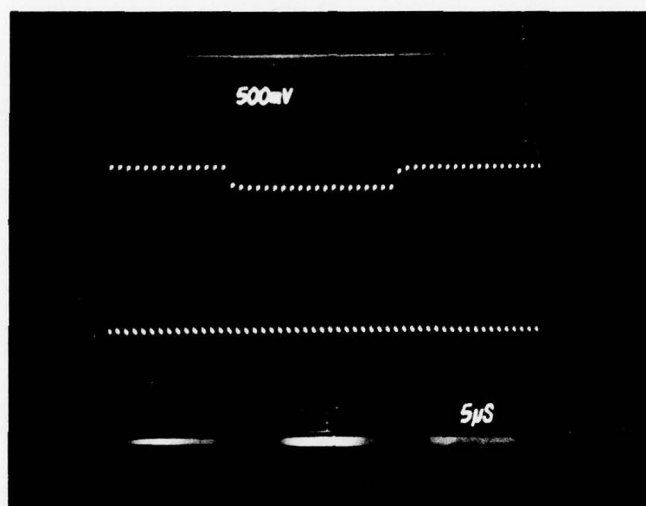
(a) Sortie au point B (figure 13)
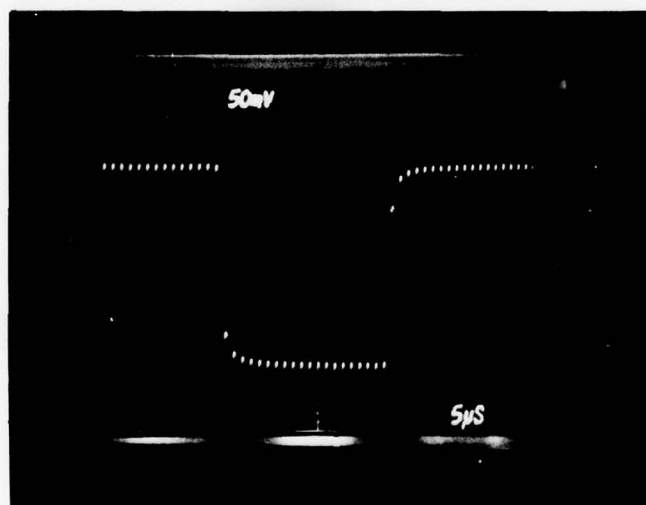
(b) Sortie amplifiée au point B

Fig.14 Résultats de l'amélioration d'efficacité sur signaux rectangulaires (suite)
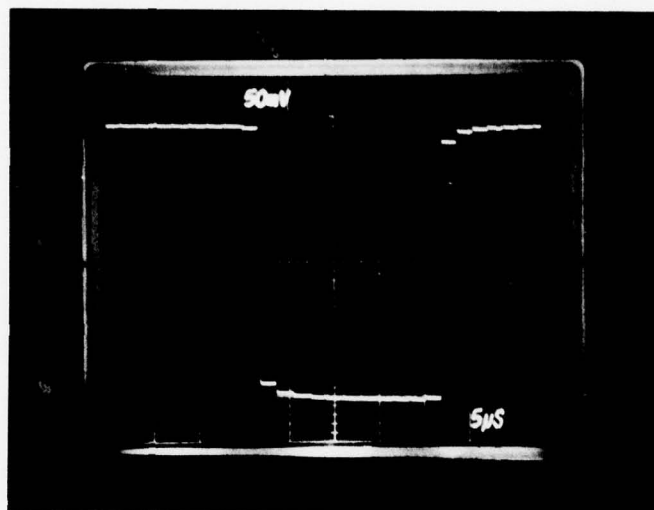
(c) Sortie au point C (figure 13)
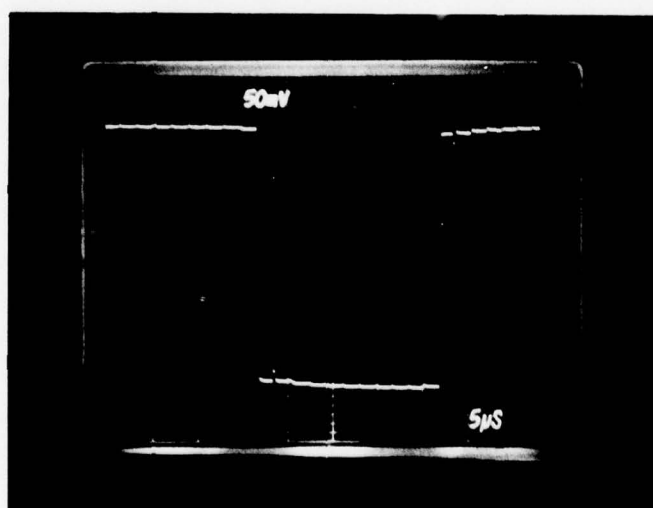
(d) Sortie au point E (figure 13)

Fig.14 Résultats de l'amélioration d'efficacité sur signaux rectangulaires (suite)
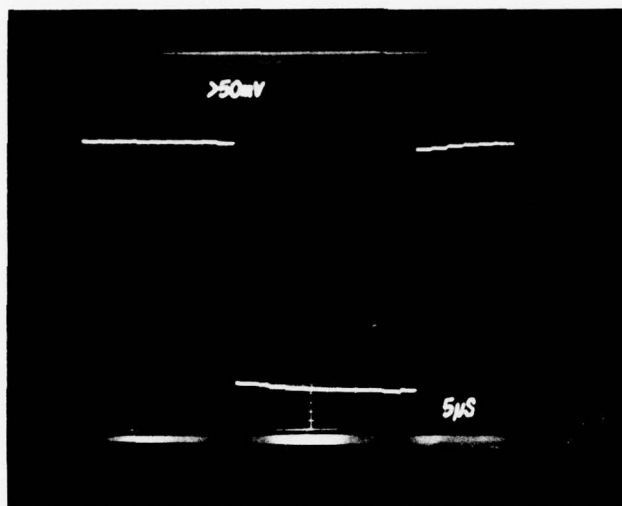
(e)   Sortie au point A′ (figure 13)
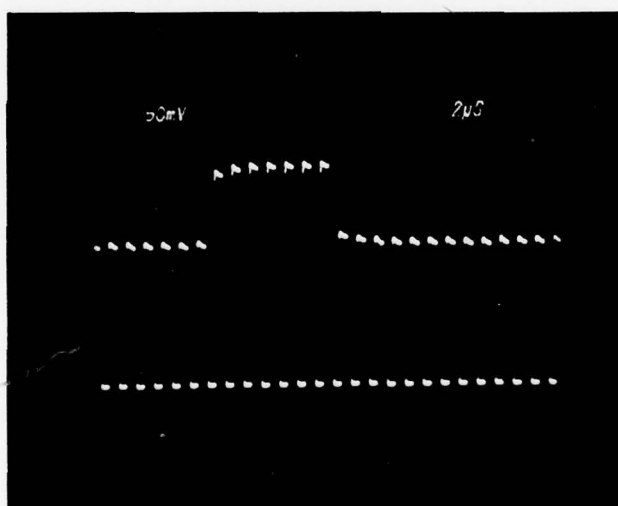
(f)   Sortie directe du CCD dans un mode différent

Fig.14   Résultats de l'amélioration d'efficacité sur signaux rectangulaires (fin)
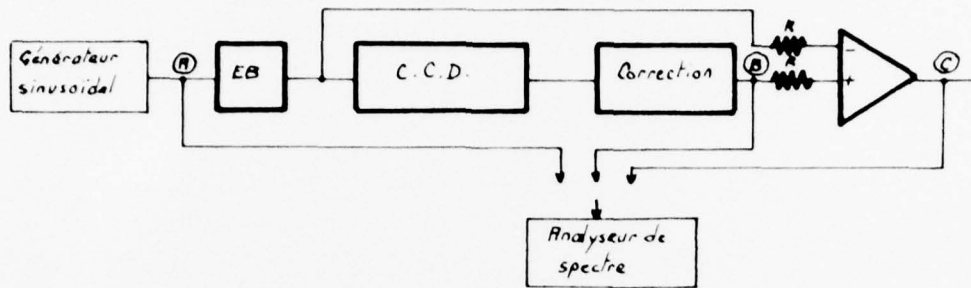
Fig.15 Filtre simple annulation
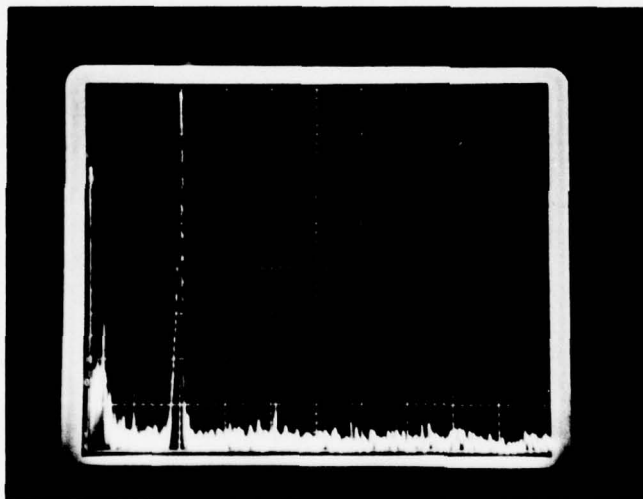
Ech
X : 10 KHz / Cm
Y : 10 db / Cm

Fig.16 Spectre de sortie du générateur
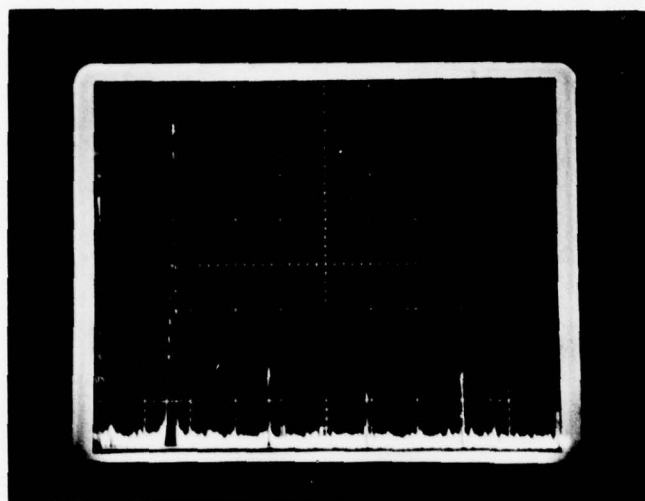
Ech.
X : 10 KHz/Cm
Y : 10 db/Cm

Fig.17   Spectre de sortie CCD après correction d'e
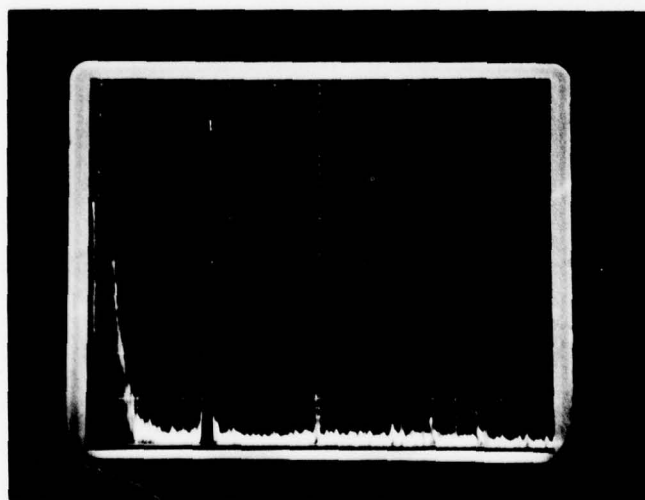
Ech:
X : 10 KHz/Cm
Y : 10 db/Cm

Fig.18   Spectre en sortie filtre pour un signal situé dans la bande passante
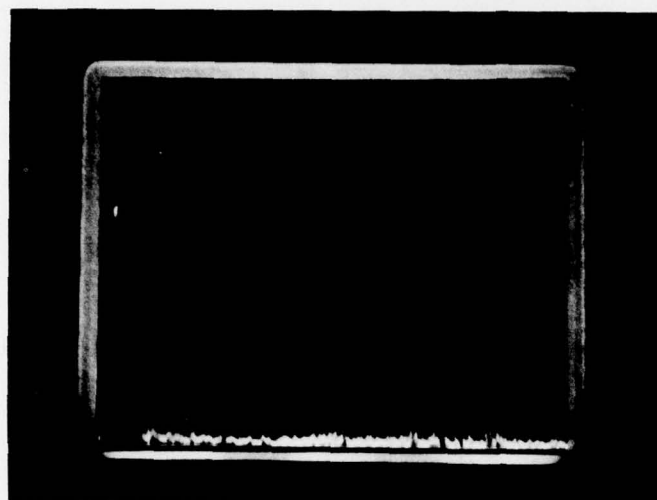
Ech
X : 10 kHz/cm
Y : 10 db/cm

Fig.19   Atténuation du filtre
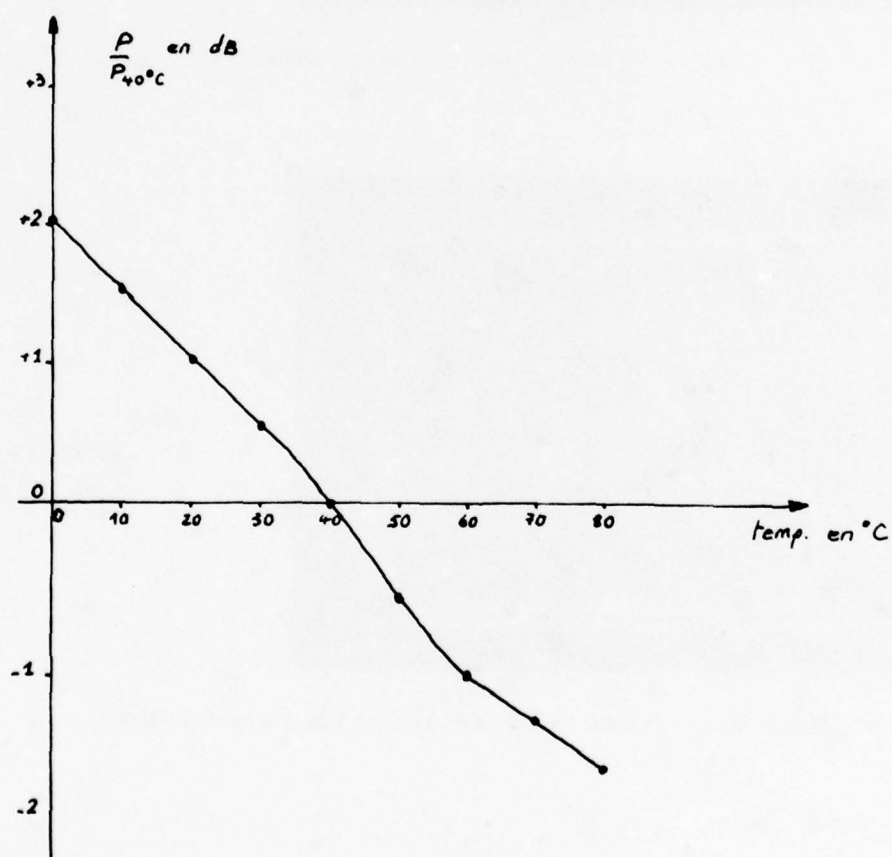
Fig.20   Influence de la température (pertes d'insertion)
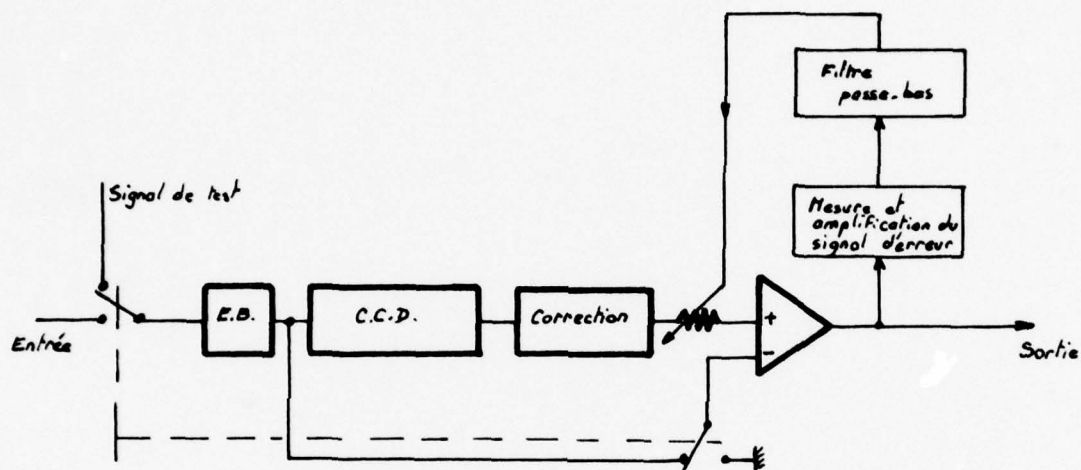
Fig.21   Principe de compensation des variations de pertes d'insertion.

A CCD DELAY LINE DOPPLER ANALYSER

J.D. Jackson, G.R. Adams, and J.S. Heeks

Standard Telecommunication Laboratories Limited

London Road, Harlow, Essex, CM17 9NA, U.K.

SUMMARY

A CCD spectrum analysing filter is reported which is compatible with the real time processing of doppler radar signals. The scheme is based on the Coherent Memory Filter (CMF) principle and has the attractive feature that the CCD function is just that of a simple delay line. Experimental results are presented for a prototype system providing a 50-point transform and performance limits, as determined by the fundamental parameters of the CCD, are discussed. In an extension of the approach it is proposed to interface the analyser with an analogue tapped delay line in order to effect both range and doppler processing.

1. INTRODUCTION

In many electronic systems useful information is carried in the frequency of the received signals and spectrum analysis becomes a basic requirement of signal processing. For optimised detection additional demands may be imposed on the processing such as the ability to detect in the presence of noise (i.e. to match filter), to fulfil the resolution capability of the system, and to provide a spectral measure in real time. Traditionally such processing has been performed with a contiguous filter bank that examines each frequency component simultaneously but for good resolution and signal enhancement requires a large number of narrow band filters. In recent years new transform techniques have emerged, the Fast Fourier Transform (FFT) increasing the speed of digital processors and the chirp Z transform (CZT) allowing compact solutions to real time processing with analogue components such as SAW and CCD. An additional analogue scheme is based on the Coherent Memory Filter (SKOLNICK, M.I.) principle which in general is capable of simultaneous range gating and doppler filtering and whose realisation in the SAW technology has been reported (JACKSON, J.D. et al., 1976). The technique is applied here to doppler-only processing and since the bandwidth required for range processing is eliminated the system is conveniently implemented with a CCD delay line.

The approach offers an interesting contrast to the CZT in that it allows a simpler CCD device in which complex transversal weighting is eliminated at the expense of increased component bandwidth. In principle both transforms require an identical time bandwidth product for the CCD element but the CMF has the advantage of being compatible with direct implementation as a doppler analyser. To meet the high performance (e.g. resolution, dynamic range) of modern radar the delay line is required to have good characteristics since signals are processed for the full effective dwell time and suffer the corresponding long time effects of the CCD. However, the functionally simple untapped delay is the most convenient CCD component with which to achieve a high level of performance.

Included in the paper are results of a prototype demonstration model and a discussion, backed by computer simulated results, of performance degradations imposed by the CCD element. Finally a scheme is outlined which allows the processing to be extended to more general signals and to facilitate range gating.

2. PRINCIPLES OF THE TRANSFORM

In common with the digital FFT and the CCD implementation of the CZT (MAYER, G.J., 1975) the transform performed by the CMF is a structuring of the Discrete Fourier Transform (DFT). This is performed within a recirculating loop in three basic operations, namely

delay, complex multiplication and summation. This process is illustrated by the schematic circuit of Fig. 1. Signal entering the input port is circulated through the delay line (T), frequency offset by δf, amplified to overcome loop losses, band limited, and cumulatively summed. In the absence of the frequency offset the loop is of course a simple delay line integrator with the well known comb filter response. Introducing an appropriate offset provides a time varying phase shift which effectively tunes the filter over the required spectrum coverage and allows real time spectrum analysis of an input signal.

To provide a better insight of the spectral analysis we can consider the circuit as a sampled data system. This is relevant in determining the output at a particular time (t) since it will consist of the input signal at (t) summed with previous inputs at multiples of the loop delay. Thus time may be referenced to the start of a delay line cycle by assigning $t = \tau + nT$, where n is an integer and $0 \leqslant \tau < T$. Synchronising the offset frequency with the loop delay, e.g. condition $\delta f = 1/T$, establishes an offset phase $2\pi t/T$ which repeats with period T and has invariance $2\pi \tau/T$. The output at (t) therefore consists of the summation of the delayed and phase shifted samples and can be expressed mathematically as

$$F(\tau) = \sum_{n=0}^{N-1} f_n \, e^{-j2\pi n \, \tau/T}$$

where f_n is the input signal and the summation is over the number of data circulations. The transformed function $F(\tau)$ appears in real time and for an input bandwidth $1/T$ the transformed frequency is proportional to time. The complex exponential represents the Fourier matrix of the DFT and in character to this transform the output of the CMF for a cw input is a sinusoid with an envelope function $\sin Nx/\sin x$ which repeats every T seconds. A complete derivation of this function has been given by Capon (CAPON, J., 1960).

For an N-sample signal the peak amplitude of the response is N times greater than the input amplitude and the frequency resolution is of order $1/TN$. Unweighted the response is $\sim \sin x/x$ and weighting techniques are usually required to reduce the sidelobes to a lower level.

3.   DOPPLER ANALYSING FILTER

3.1.   Implementation Considerations

A doppler analysing filter based on the CMF requires a delay line with the following parameters. Its delay must be equal to the reciprocal of the doppler bandwidth (equal to the PRI or code repeat period in a pulse radar) and its frequency coverage equal to the product of the doppler bandwidth and the number of doppler cells. For a typical radar this means a delay of $\sim$ 100 μs and a bandwidth of $\sim$ 1 MHz. A CCD analogue delay adequately meets such requirements and offers the advantage that it can be precisely controlled and locked to the offset frequency.

In common with any close-loop recirculating system the design is complicated by the cumulative effect of component imperfections and the performance becomes sensitive to signal transfer characteristics such as amplitude and phase distortion, and time and frequency spurii. Concerning the CCD implementation, this will influence the quality of the system to an extent determined by such parameters as charge transfer efficiency, input/output nonlinearity and dark current. Charge transfer inefficiency manifests itself as a frequency dependent amplitude attenuation and a frequency dependent phase shift (VANSTONE, G.F., et al., 1974). The amplitude attenuation leads to a variation of loop gain and with additional loss from peripheral circuits (e.g. sample and hold) the gain roll-off can be up to 3 dB over the loop bandwidth. The loss can be reduced with a correction amplifier although the effect of any residual gain variation can be

cumulative and can rapidly attenuate the frequency samples which appear at the output. For example, a loop which introduces a linear amplitude taper of 0.5 dB over its bandwidth will produce a maximum attenuation of the output samples of 28 dB for a loop which supports 100 circulations. In practice this can be equalised by amplitude weighting the input samples and this operation can be performed in conjunction with prescribed weighting for sidelobe reduction (see section 3.2). Computer analysis of random amplitude errors has shown that the weighting coefficients should be implemented with 1% accuracy in order to achieve -40 dB sidelobe response.

The effect of phase errors in the loop is to shift the doppler indication in time and to degrade the peak to sidelobe response. For minimal degradation we have calculated that the deviation from linear phase must be no greater than 1 degree although any greater non-linearity can in principle be corrected by phase weighting the input signal. To assess the phase error contribution from the CCD component, the doppler analyser has been computer modelled. Fig. 2(a) shows the response for an ideal CCD delay. The input bandwidth is 10 kHz and the input doppler is 5 kHz; there are 50 circulations in the loop and the signal is weighted for -40 dB sidelobe response. Fig. 2(b) shows the effect of a transfer efficiency of 95% over the complete delay. For a typical CCD having 500 stages of delay and 2 phases per stage this represents an inefficiency (ϵ) of 5×10^{-4} per transfer which is an order of magnitude worse than the best devices currently available. The inefficiency produces a linear phase shift which can be accommodated in the loop delay, and a phase non-linearity of ± 0.8 degree over the loop bandwidth. The effect of the latter is to perturb the base of the main lobe with a close-in sidelobe degradation of approximately 6 dB. The effect of synchronisation inaccuracy between the offset frequency and loop delay also manifests itself as a phase error and computer results indicate that synchronism to within 1 part in 10^5 is required. The stability that this implies of system frequencies imposes no problem in practice.

The non-linearity of the CCD causes harmonic distortion and introduces frequency domain spurii. Typically, the non-linearity is of order 3% which limits dynamic range (i.e. signal to spurii) to $\sim$ 30 dB. This can be improved by using a linearisation technique (MACLENNAN, D.J., et al., 1975 ; McCAUGHAN, D.V., et al., 1976), allowing a non-linearity of better than 1% to be achieved and extending the useful dynamic range to 40 dB. Other parameters of the CCD (e.g. dark current and signal to noise ratio) would allow 50-60 dB dynamic range.

3.2. Experimental Prototype

The following radar system parameters have been assumed for the purpose of demonstrating the system concept.

signal repetition interval	=	100 µs
doppler bandwidth	=	10 kHz
number of doppler cells	=	50
time on target	=	5 ms

The CCD analogue shift register is clocked to provide 100 µs of delay and the offset frequency set to 10 kHz. With 50 circulations in the loop the input bandwidth of 10 kHz is resolved into 50 doppler cells each $\sim$ 200 Hz wide. The total processing time is then 5 ms which, for maximum resolution, would be equal to the time on target. CCD's are usually operated with a baseband signal input but, given an adequate component bandwidth, it is possible to avoid the use of I and Q channels by operation at a minimum value IF. This approach has been adopted in the present system, a convenient carrier frequency being about 0.6 MHz.

The circuit implementation of the doppler analyser is shown in Fig. 3. The delay element is a Fairchild CCD-321 with one of two independent 455 stage lines being clocked at 4.55 MHz to establish a 100 µs delay. A sample and hold feature within the device reconstructs the analogue signal. Following the CCD a 7-pole Butterworth low-pass

filter further reduces the clock breakthrough and limits the aliased frequencies. This filter has a corner at 1.1 MHz which also limits the frequency excursion in the loop. In order to maintain low spurious levels in the signal band the frequency offset is provided by a single sideband mixer which works via double frequency conversion and sideband filtering. Each mixer L.O. and the CCD drive are phase locked to a master clock which ensures exact synchronisation between offset frequency and time delay.

Weighting can be implemented by frequency filters in the loop or on the output, or by time weighting the input signal. The latter is preferred on the basis of ease of implementation together with the ability to correct any non-optimum weighting imposed by say the antenna or loop response. The required circuit is simply an amplitude modulator with a preset pattern of independent levels which are cycled through in synchronism with the delay period. An output gate ensures the correct observation after each integration period.

Some examples of the operation of the unit are given by the spectra shown in Figs. 4 and 5. The response to a c.w. input is shown in Fig. 4; the doppler is 7 kHz and the resolution corresponding to 50 circulations is 200 Hz at the -3 dB points. In character with the transform the spectrum is seen to repeat at intervals of 100 µs. Fig. 5 shows the response to an FM input signal having a carrier frequency of 0.605 MHz; the deviation is 2.5 kHz, and the modulation 2.2 kHz. The responses are essentially unweighted apart from incidental weighting imposed by the loop filter which has a roll off at 1.1 MHz. Accordingly the peak sidelobe level is around -13.5 dB. Further, no attempt has been made to linearise the phase response of the loop filters and evidence of phase distortion is observed in the asymmetry of the sidelobe response. Improvements to fully optimise the operation are in hand.

4.   PULSE-DOPPLER PROCESSOR

The doppler analyser described in section 3 has a processing capability limited to input signals of modulation bandwidth less than the doppler bandwidth; this is characteristic of the DFT process. This section describes a system for extending the features of the analyser to handle signals of more general format and increased modulation bandwidth while being able to access range information via matched filtering.

A schematic of the processor is shown in Fig. 6 in which the first stage of matched filtering is performed with a tapped delay line correlator (MENAGER, O., et al., 1975; DARBY, B.J., et al., 1975). In radar operation the transmitted pulse is delayed by multiples of the pulse width or code rate and correlation is achieved by multiplying the delayed replicas by the target return. Each parallel channel defines a range cell and full coverage is achieved when the total delay is equal to the PRI or code repeat period. When the target return has a doppler shift a sampled doppler appears at the output of the multiplier and a low pass filter passes the doppler frequency. Simultaneous range and doppler processing is achieved by connecting a doppler analyser to each parallel channel although less complex system using one analyser are possible with reduced processing capacity.

The tapped delay line may be implemented in SAW or charge transfer device (CTD) technology. With the latter the only difference is that the output from the multiplier appears as a discrete time correlation. SAW tapped delay lines provide the bandwidth for modest (20 MHz) data rates and will be useful in radar applications that require high resolution ranging over short (e.g. 1-2 km) ranges or precision tracking over a few range cells. A CTD device will provide a longer delay but lower bandwidth capability than SAW and will be applicable to low range resolution search radars and low data rate communications. The inherent capability of the processor to allow simultaneous search in velocity and to operate with diverse waveforms will be of particular relevance to spread spectrum systems and in ECM environments.

Available components are being assembled to establish the feasibility of the processor. A commercially available bucket brigade device (Reticon TAD-32) with 32 taps is being assessed for handling low bandwidth (e.g. 1 MHz) pulse waveforms while a 31-tap SAW filter will be used to demonstrate operation with 20 Mbit/s code rates.

5.   CONCLUSIONS

A doppler analysing filter has been demonstrated which uses a CCD in a particularly simple architecture when compared to other processors. It performs spectrum analysis by effecting the exact DFT and accordingly operates in a continuous manner - unlike the standard form of CZT which requires a second interlacing processor.

Implementation studies of the system have indicated that the quality and accuracy required both of the CCD and the peripheral electronics are within the capabilities of the present technology and transform lengths of 50-100 with approximately 35 dB spectral sidelobes are predictable. The experiments so far have shown that a resolution equivalent to a 50 point transform is achievable but in terms of spectral sidelobes and dynamic range the system is below optimum and further design and circuit implementation improvements are required.

6.   REFERENCES

CAPON, J., 1960, "On the Properties of an Active Time-Variable Network:The Coherent Memory Filter", Proc. of the Symp. on Active Networks and Feedback Systems, Polytechnic Inst. of Brooklyn.

DARBY, B.J. and MAINS, J.D., 1975, "The Tapped Delay Line Active Correlator:A Neglected SAW Device", 1975 Ultrasonics Symp. Proc. IEEE Cat. No. 75 CHO 994-4SU, p. 193-196.

JACKSON, J.D., JONES, E. and HEEKS, J.S., "Pulse Doppler Processing with a SAW Delay Line", 1976 Ultrasonics Symp. Proc. IEEE Cat. No. 76 CH 1120-5SU, p. 505.

MACLENNAN, D.J. and MAVOR, J., 1975, "Novel Technique for the Linearisation of Charge-coupled Devices", Elect. Lett., 11, (10), p. 222-223.

McCAUGHAN, D.V. and HARP, J.G., 1976, "Phase-referred Input:A Simple New Linear CCD Input Method", Elect. Lett., 12, (25), p. 682.

MAYER, G.J., December 1975, "The Chirp Z-Transform - A CCD Implementation", RCA Review, 36, p. 759-773.

MENAGER, O. and DESORMIERE, B., 1975, "Sampling Correlator using Acoustic Surface Waves", App. Phys. Lett., 27, (1), p. 1.

SKOLNICK, M.I., (ed), Radar Handbook, McGraw-Hill, p. 17-55.

VANSTONE, G.F., ROBERTS, J.B.G., and LONG, A.E., 1974, "The measurement of Charge Residual for CCD Transfer using Impulse and Frequency Responses", Solid-State Electronics, Vol. 17, p. 889-895.
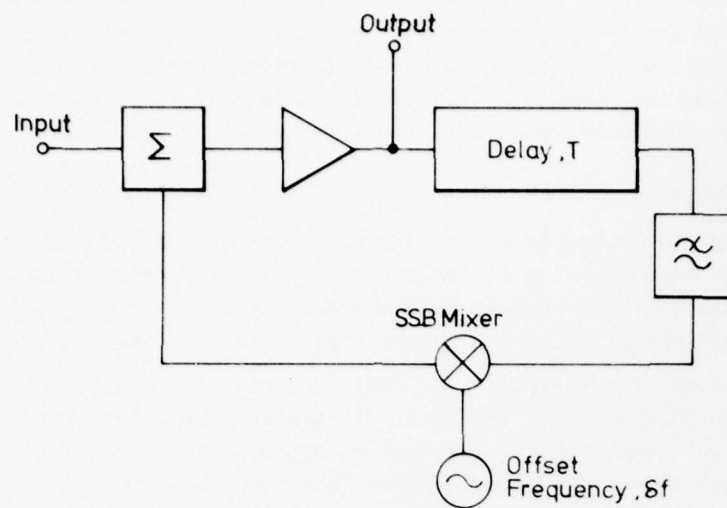
Fig. 1   Principle of the Coherent Memory Filter.
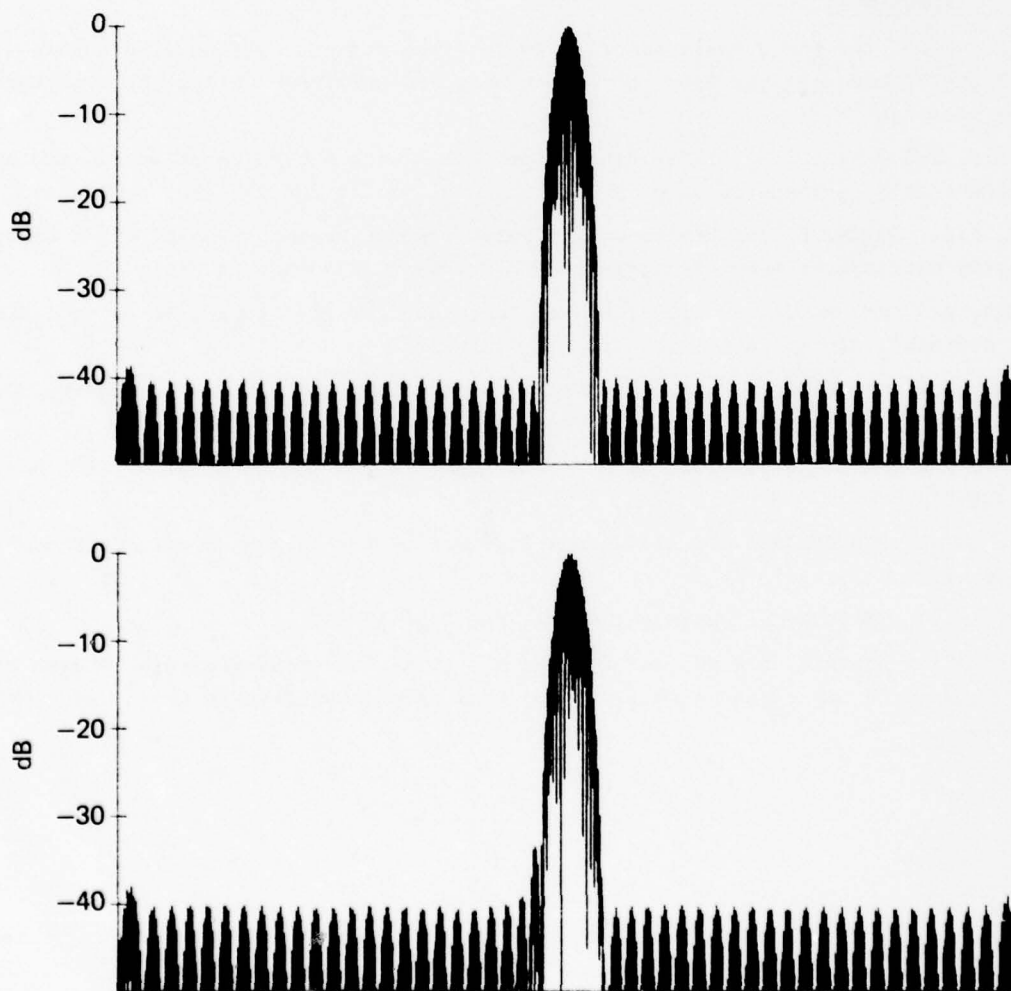
Fig. 2   Computer simulation of 50-point CMF transform
           for

        (a)   Ideal delay line (upper trace),

and (b)   CCD Delay line with transfer inefficiency
           $\varepsilon = 5 \times 10^{-4}$ (lower trace).

Output
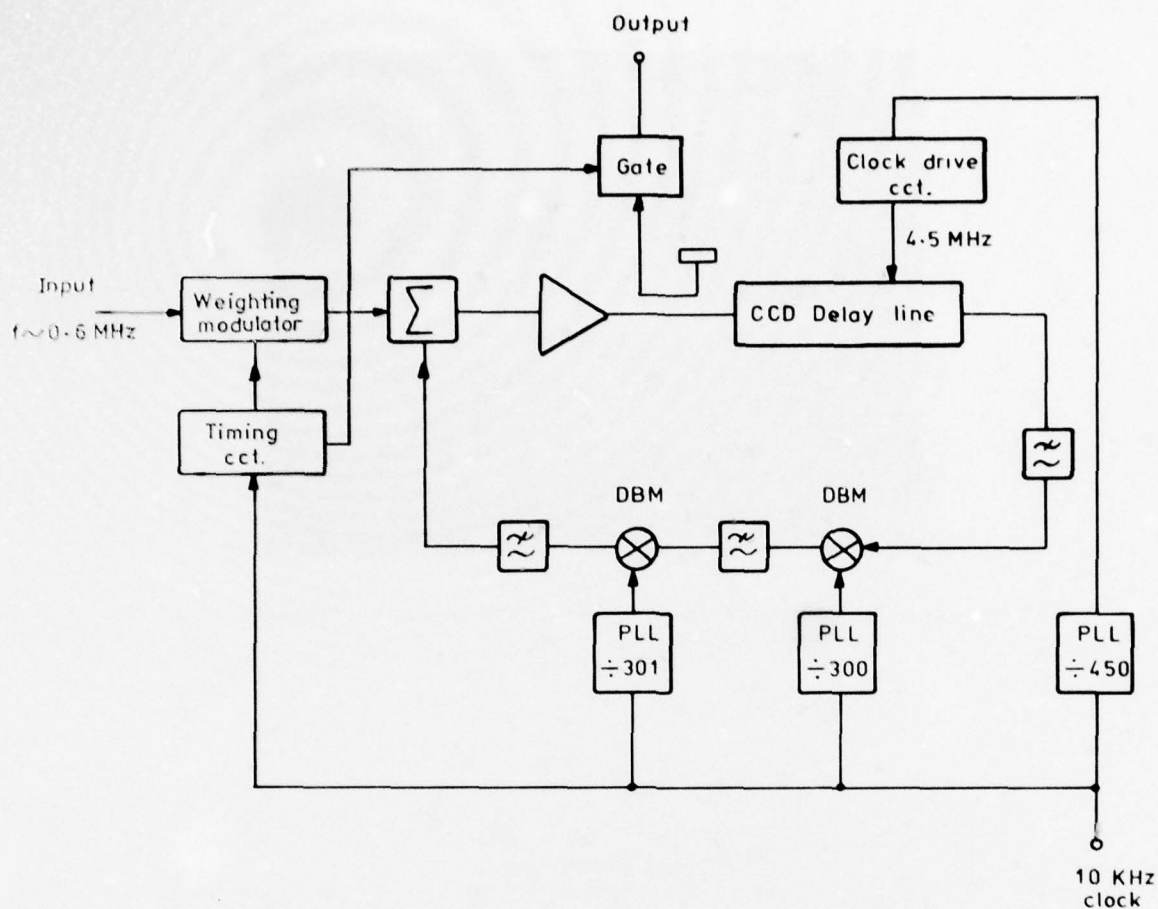
Fig. 3   Schematic of CCD Doppler Analyser.
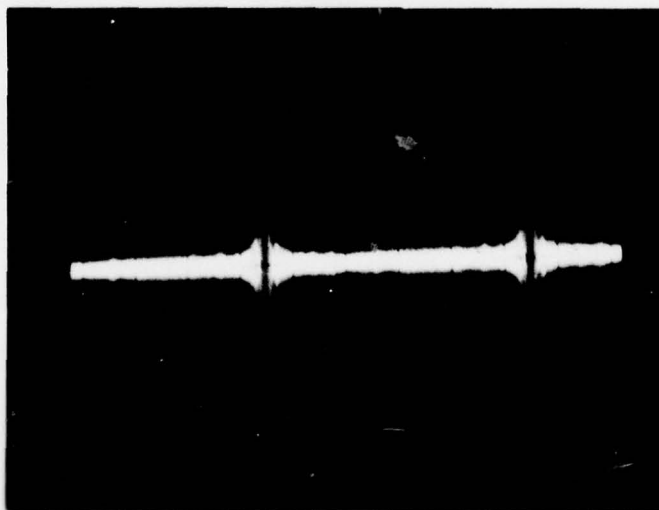
Fig. 4   Response of analyser to C.W. signal; horizontal
scale 2 kHz/div.
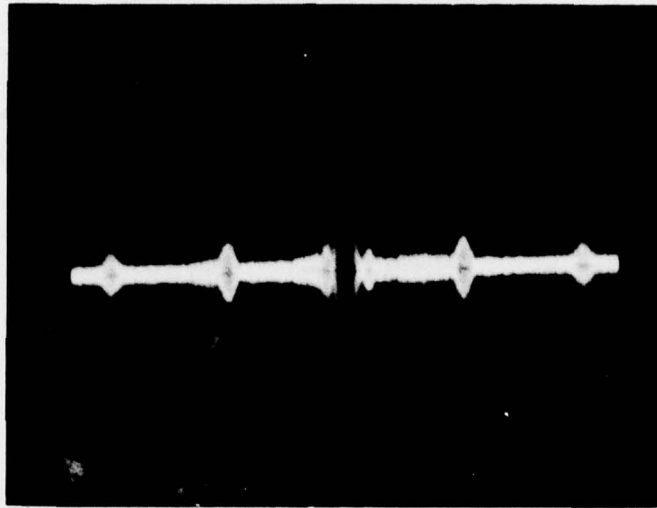
Fig. 5   Response of analyser to FM signal; horizontal
scale 1 kHz/div.
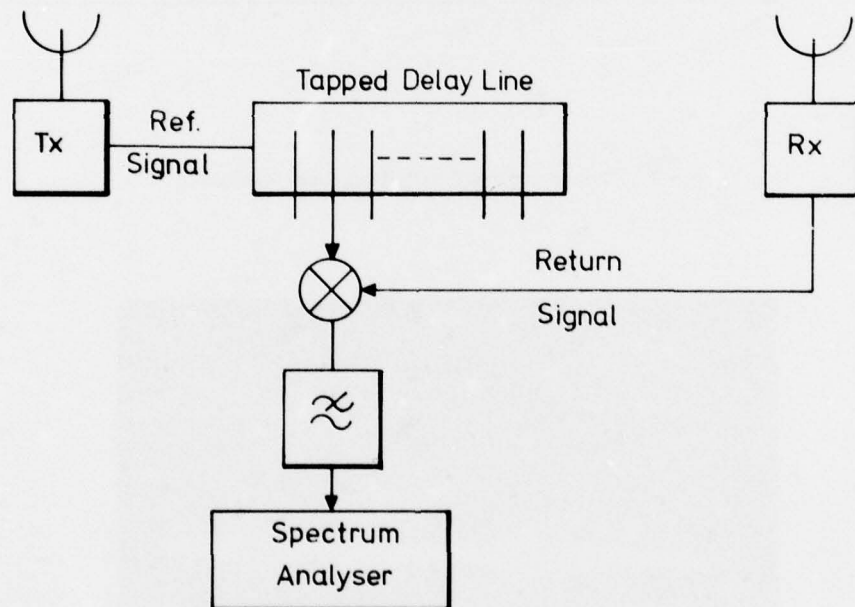
Fig. 6   Pulse Doppler Processor.

COMBINED ACQUISITION AND FINE SYNCHRONIZATION SYSTEM FOR SPREAD SPECTRUM RECEIVERS USING A TAPPED DELAY LINE CORRELATOR

by

W.P.Baier and M.Pandit
University of Kaiserslautern, P.O.Box 3049, D-6750 Kaiserslautern

and

H.Grammüller
Siemens AG, Long Range Communication Division, Joint Divisional Development,
P.O.Box 700070, D-8000 Munich 70

SUMMARY

In direct sequencing spread spectrum communications systems a local PN generator must be synchronized with the PN code contained in the incoming signal at the receiver. To acquire synchronism and maintain it (i.e. for "acquisition" and "tracking") one incurs considerable expense for the receiver. It is known that the acquisition time and the complexity of the acquisition circuitry can be considerably reduced by employing a tapped delay line as a matched filter. The paper presents a method of using a tapped delay line correlator not only for acquisition but also for tracking. One employs at the outset PN clock generators of high inherent stability and low relative drift at the transmitter and receiver. The correlation peaks at the output of the tapped delay line correlator are used to set the receiver PN generator in phase with the received PN code. During the intervals of time between the correlation peaks the local PN generator runs freely. A system - including the hardware realization - based on this principle is presented. The salient theoretical results have been checked by measurements.

1. INTRODUCTION

Spread spectrum communications systems, in which binary PN sequences are employed for bandspreading, are frequently used to obtain jam resistance. Systems of this type have the additional advantage that they also offer the possibility of multiple access, message privacy and range measurement.

Bandspreading by means of binary PN sequences can be produced by frequency hopping, time hopping or direct sequencing (CAHN, C.R. et al.,1973; DIXON, R.C.,1976). This paper examines communication systems with direct sequencing of the basic design shown in Fig.1. The message which may be either in digital or analog form is first modulated onto a carrier by a conventional message modulator in the transmitter. The relatively narrow-band signal thus obtained is spectrally spread by subsequent rapid pseudo-random phase shift modulation in a biphase modulator and then transmitted. In the receiver, at first the pseudo-random phase reversals are removed, once again in a biphase modulator. The signal spectrally compressed in this way is passed through a bandpass filter whose band-width is identical with that of the compressed signal and fed to the message demodulator. The demodulator output yields the message.

In the biphase modulator of the receiver, the spectrum of narrow-band interfering signals is spread and that of wideband interfering signals remains virtually unaffected whereas the spectrum of the desired signal is compressed. The bandpass filter at the receiver thus prevents a substantial proportion of the interfering power from reaching the message demodulator.

The basic design of the spread spectrum communication system shown in Fig.1 differs from a "conventional" communication system only in that it has a few additional circuits: Two biphase modulators, a bandpass filter and two binary PN generators. However, circuitry additional to that shown in the receiver in Fig.1 is necessary to generate the local PN sequence which must be identical and synchronous with the PN sequence contained in the received signal. Identity of the codes can easily be obtained by using identical PN generators in the transmitter and receiver. However, it is more difficult to synchronize the local PN generator. The problem of synchronization can be resolved into two subsidiary problems:

    1) Obtaining initial synchronization (acquisition),
    2) Maintaining synchronism (tracking).

The circuits necessary for synchronization render the receivers of spread spectrum communication systems more complicated and hence larger, heavier and more expensive than the receivers of conventional communication systems.

The two most widely used synchronization techniques, which appear today in various forms, are shown schematically in Figs.2 and 3. The technique shown in Fig.2 is based on the principle of active correlation (SAGE, G.F.,1964; WARD, R.B.,1965). For acquisition, the local PN generator is clocked in such a way that its PN sequence is shifted continuously or in discrete steps relative to the sequence contained in the received signal. Thereby occasional instants of synchronism are produced which are detected by the acquisition circuit. The latter activates the control loop provided for tracking, which is usually a delay locked loop (GILL, W.J.,1966).

The technique shown in Fig.3 is based on the principle of passive correlation in which the acquisition circuit has the task of detecting the instant at which the received PN sequence attains a predetermined state. At this moment the local PN generator is started from the corresponding initial state. At the same instant the delay locked loop, also provided for tracking in this system, is closed. A matched filter, which is matched to the predetermined section of the received signal is a basic component of the acquisition circuit in the case of passive correlation (CAHN, C.R. et al.,1973; DIXON, R.C.,1976). The output signal of the matched filter has periodic amplitude peaks which can be

detected by a suitable threshold circuit. The matched filter can be implemented with the aid of tapped delay lines. The most suitable devices for use in tapped delay line correlators are delay lines for acoustic surface waves (SAW) or shift registers for analog signals in the form of charge coupled devices (CCD) (BELL, D.T. et al.,1973; HAYS, R.M. and HARTMANN, C.S.,1976; BUSS, B.D. et al.,1973; DAVIES, N.G. and WOHLBERG, D.B.,1976).

A synchronization circuit of the type shown in Fig.3 with a matched filter is generally simpler and cheaper to implement than a synchronization circuit as shown in Fig.2. This reduction in cost applies however, only to the acquisition circuit. The considerable complexity of the tracking circuit, i.e. the delay locked loop, is the same in both synchronization methods. It would be advantageous if tapped delay lines could also be used to simplify the tracking circuit. The authors have set themselves the task of designing a simple spread spectrum receiver in which one and the same tapped delay line is used both for acquisition and for tracking. The design of a receiver of this type, in which the peaks of the tapped delay line correlator periodically set - and thus synchronize - the otherwise free-running local PN generator is dealt with in the following.

The paper is devided into four sections. In Section 2 the operation of a tapped delay line correlator for signals with PN biphase modulation is briefly recapitulated. Section 3 explains the proposed synchronization techniques. In this section the design of a spread spectrum receiver in which acquisition and tracking is performed by means of the same tapped delay line correlator is described. Section 4 reports on practical work which corroborates the theory developed in the previous sections.

2. THE TAPPED DELAY LINE USED AS A MATCHED FILTER FOR SIGNALS WITH PSEUDO-RANDOM BIPHASE MODULATION

Fig.4 shows a schematic diagram of a tapped delay line, including the additional circuitry for use as a matched filter in spread spectrum receivers. The input signal $s_1(t)$ of the delay line - the uncorrupted received signal in the interference-free case first considered here - has a carrier angular frequency of ω_0 and an amplitude A. Its phase is reversed at the rate of the binary PN sequence $p(t) \in \{-1,1\}$ with the clock frequency $f_C = 1/T_C$. The sequence $p(t)$ is usually a m-sequence but it can also be any periodic binary sequence (DIXON, R.C.,1976) whose autocorrelation function $R(\tau)$ at $|\tau| \geq T_C$ is very low. If the narrow-band message modulation is initially ignored, we obtain

$$s_1(t) = p(t) \cdot A \cdot \cos(\omega_0 t). \tag{1}$$

The chip period T_C is selected in such a way that it is a whole-numbered multiple of the carrier period. The delay line has k taps spaced at time intervals of T_C. In accordance with a section of sequence $p(t)$ with a length of k chips, k_H of the k output signals of the delay line are fed directly to a adding network and the remainder k_L (=k - k_H) after a 180° phase reversal. The signal at the output of the adding network is denoted by $s_2(t)$. Ignoring the insertion loss of the delay line, the amplitude of $s_2(t)$ is in the order of A for most of the time since approximately k/2 signals having a phase of 0° and approximately k/2 signals having a phase of 180° are added. Only when the actual phase reversals of the signal $s_1(t)$ match those reversals programmed into the delay line has $s_2(t)$ the amplitude kA since k signals having the same phase are added. By rectifying the signal $s_2(t)$, one obtains the signal $s_3(t)$. The signal $s_3(t)$ attains a maximum every time the amplitude of $s_2(t)$ equals kA, and has much lower values at other times. The peaks of $s_3(t)$ are detected with a threshold circuit. The section of the PN sequence which is "programmed" into the tapped delay line must be chosen so that the main correlation peak exceeds parasitic correlation peaks considerably. In practical systems, a narrow-band message modulation is also contained in $s_1(t)$. This impairs the operation of the correlator. The impairment can be made negligibly small as long as the bandwidth of the message modulation is not too large. If an interfering signal is present, the signal-to-noise ratio at the output of the delay line is higher by about a factor k as compared to the input signal-to-noise ratio (BELL, D.T. et al.,1973). The non-detection probability and false alarm rate with which the peaks of $s_3(t)$ are detected decrease therefore as k increases. The circuit consisting of a tapped delay line matched filter, rectifier and threshold detector shown in Fig.4 is referred to in the following as a "tapped delay line correlator".

The peaks which - due to the pseudo-randomness of $p(t)$ - periodically occur in $s_3(t)$ can be used, as already explained in Section 1., in a passive correlation acquisition circuit as shown in Fig.3. The following section describes the principle of a system in which the peaks at the output of the tapped delay line correlator are not only used for acquisition purposes but also for tracking.

3. SPREAD SPECTRUM RECEIVER WITH TAPPED DELAY LINE CORRELATOR FOR ACQUISITION AND TRACKING

If we were to assume that PN generators with absolutely stable identical clock frequencies could be employed in the transmitter and the receiver and that no frequency shifts due to Doppler effect arise, it would be possible to synchronize the otherwise free-running local PN generator by correctly setting and starting it once only. Unfortunately this simple synchronization method is unrealistic. In practice, even the most stable clock oscillators have a certain drift, and in many applications there is relative movement between the transmitter and receiver and hence a frequency shift due to Doppler effect. It is however possible to modify the synchronization scheme based on "setting and starting" so that it is practically realizable. For this purpose one must ensure that the local PN generator is set and started correctly afresh whenever the local PN sequence has shifted too much relative to the received sequence, i.e. whenever the "epoch error" has become too large. Repeated setting and starting in this way solves the problems of acquisition and tracking simultaneously.

The periodic peaks of the output signal $s_3(t)$ of the tapped delay line correlator shown in Fig.4 mark a certain state of the received PN sequence. The signal $s_3(t)$ is therefore suitable as a synchronization pulse for repeated setting and starting of the local PN generator. A receiver design in which synchronization is based on this principle is shown in Fig.5. The PN generator consisting of clock oscillator and shift register is periodically set to the instantaneously correct state and started by the tapped delay line correlator. This process of setting and starting is necessary only if an epoch error has occured since the previous synchronization. If the PN clock oscillators are sufficiently stable, the loss of individual synchronization pulses does not result in a loss of synchronization whereas any false alarm breaks the synchronism until the next correct synchronization pulse appears. A comparison of Fig.5 with Figs.2 and 3 shows the simplification offered by the system in Fig.5. As a further advantage the detector lock-on circuit required in the designs with delay locked loop can be dispensed with.

For the receiver design shown in Fig.5 it is essential that the epoch error of the local PN sequence arising between two neighbouring synchronization pulses is not too large and that the setting and starting is sufficiently accurate. These requirements have to be met with by suitably chosing the system parameters. The epoch error which builds up between two successive synchronization pulses is mainly given by the relative drift of the two clock oscillators and by Doppler effect. The setting and starting accuracy is impaired by interference at the receiver input. The interference causes a certain time error of the synchronization pulses. In addition, interference gives rise to false synchronization pulses and prevents the detection of some of the correct synchronization pulses. The rate of these false alarms and losses as well as the time inaccuracy of the detected synchronization pulses become greater as the signal-to-noise ratio γ_e at the receiver input becomes smaller and as the deviation of the received signal from (1) due to the narrowband message modulation increases.

3.1. QUANTITATIVE EXAMINATION OF THE PROPOSED SYSTEM

The performance of the system shown in Fig.5 is examined quantitatively in this section. Analog frequency modulation is employed as message modulation. In this case the message modulator and demodulator become especially simple. The interference signal is assumed to consist of Gaussian noise which is "white" in the transmission bandwidth B_{RF}. The parameters which appear in the investigation to follow are:

Ratio B_{RF}/B_b of the available transmission bandwidth B_{RF} to the bandwidth B_b of the message to be transmitted. In order to achieve a marked degree of interference suppression, B_{RF}/B_b must be considerably greater than 1.

Length L of the PN sequence in chips.

Modulation index β of the frequency modulation (β is the quotient of the peak frequency deviation and message bandwidth B_b).

Number of taps k of the tapped delay line.

Signal-to-noise ratio γ_e at the receiver input.

Number m of consecutive time intervals LT_c during which the tolerable epoch error (e.g. $0.1T_c$) is not exceeded if setting and starting is carried out correctly at the beginning of the first of these intervals and if there are no subsequent synchronization pulses. m depends on the stability of the PN clock oscillators and on Doppler frequency shift.

With the given transmission bandwidth B_{RF} the clock frequency f_c is usually so chosen that

$$f_c \approx B_{RF}/2 \tag{2}$$

holds. This will be adopted in the following though clock frequencies different from $B_{RF}/2$ might give better jam-resistance (see also BAIER, W.P. and MEFFERT, K.,1977).

The aim of the quantitative investigation is to obtain an insight into the jam resistance and synchronization performance of the system. As a measure for the synchronization performance, the out-of-lock time t_{out} is introduced. t_{out}, expressed as a fraction of the overall transmission period, indicates the period of time during which the epoch error is greater than the permissible maximum value of, say $0.1T_c$. The time error of the instants at which the treshold is exceeded can be given by their standard deviation τ_{rms} from the ideal instant. A measure for the jam resistance is the signal-to-noise ratio γ_a which can be obtained at the output of the message demodulator. In the following, the quantities t_{out}, τ_{rms} and γ_a are calculated as functions of the given quantities B_{RF}/B_b, γ_e, k and m.

3.1.1. OUT-OF-LOCK TIME T_{OUT}

If t_{out} is calculated as a function of P_d, n_{fa} and m, one obtains the expression

$$t_{out} = (1-P_d)^m + \left[1-(1-P_d)^m\right] n_{fa}LT_c/2. \tag{3}$$

The greater the value of P_d and the smaller the value of n_{fa}, the smaller is t_{out}. The following equations are obtained for P_d and n_{fa} with the normalized threshold value b' of the threshold detector in Fig.4:

$$P_d = Q(\sqrt{2k}\,\gamma_e,b'), \tag{4}$$

$$n_{fa} \, LT_c = \sqrt{\frac{2\pi}{3}} \; \frac{b'}{\sqrt{1+\gamma_e}} \cdot L \cdot \exp\left[\frac{-b'^2}{2(1+\gamma_e)}\right] \quad . \tag{5}$$

Equations (3), (4) and (5) are derived in Appendix I. The function Q in equation (4) is the Marcum-Q-function. It is assumed in (4) and (5) that the tapped delay line improves the signal-to-noise ratio by a factor k. If γ_e, L and T_c are fixed, the quantities P_d and n_{fa} in (3) are determined by the normalized threshold value b'. Both P_d and n_{fa} increase as b' decreases. A decrease of the threshold value b' therefore has a favourable effect on t_{out} in one respect, but an unfavourable one in another; it is to be expected that there is an optimum threshold value b' at which t_{out} is a minimum. The authors have evaluated expressions (3), (4) and (5) to obtain the optimum b' and the minimum out-of-lock time as a function of γ_e, k and m for the length L = 1023 of the PN sequences. Fig.6 shows the results. The minimum t_{out} is shown as a function of m with γ_e and k as parameters.

It must be emphasized that the applicability of the presented synchronization scheme is not confined to systems which use frequency modulation as message modulation. Another suitable form of message modulation is for example PSK. In this case the message modulation and demodulation equipment becomes more complicated. However, the effects of non-constant carrier frequency which are considered in Section 3.1.3. can be avoided.

3.1.2. TIME ERROR OF THE SETTING SIGNAL

If f_c is chosen according to (2), the envelope of the correlation peak at the output of the tapped delay line can be approximated as depicted in Fig.7. In this case the standard deviation τ_{rms} of the instant at which the threshold is exceeded, normalized to T_c, is approximately given by

$$\frac{\tau_{rms}}{T_c} \approx \frac{1}{\sqrt{8k\gamma_e}} \quad . \tag{6}$$

Equation (6) is derived in Appendix II. If the envelope is assumed to have the trapezoidal form shown in Fig.7, τ_{rms} does not depend on the threshold value of the detection circuit. The slope of the envelope is not constant in practice. Measurements were carried out using the circuit shown in Fig.8, in order to check (6). The rectifier was an envelope detector. The measured τ_{rms}/T_c is shown in Fig.9 as a function of the normalized threshold value b'. It can be seen that the measured minimum values in Fig.9 agree closely with the values calculated with the aid of (6). As an example, for $k\gamma_e$ = 100 we obtain from (6) the value τ_{rms}/T_c = 0.035. The measured minimum value is τ_{rms}/T_c = 0.037.

3.1.3. SIGNAL-TO-NOISE RATIO γ_a AT THE OUTPUT OF THE MESSAGE DEMODULATOR

If analog frequency modulation is used for the message, the relationship (BAIER, W.P., 1975)

$$\gamma_a = 1.5 \cdot \frac{B_{RF}}{B_b} \, \beta^2 \, \gamma_e \tag{7}$$

holds between the signal-to-nois ratio γ_e at the receiver input and the signal-to-noise ratio γ_a at the output of the message demodulator. Equation (7) is derived in Appendix III.

Since frequency modulation is used, the carrier frequency deviates from the ideal value $\omega_o/(2\pi) = f_o$ by a maximum of βB_b. In the worst case the deviation reaches this maximum at all the taps simultaneously. As a consequence, the signal-to-noise ratio improvement due to the tapped delay line is no longer k but approximately

$$g = k \cdot \text{sinc}^2(2\pi\beta \frac{B_b}{B_{RF}} \cdot k). \tag{8}$$

A reasonable value for the argument of the sinc^2-function is

$$2\pi\beta \frac{B_b}{B_{RF}} \cdot k = 1 \quad . \tag{9}$$

This corresponds to a degradation of 1.6 dB. This case will be considered in the following. The maximum permissible modulation index is then

$$\beta_{max} = \frac{1}{2\pi} \cdot \frac{1}{k} \cdot \frac{B_{RF}}{B_b} \quad . \tag{10}$$

By substituting (10) in (7) one obtains

$$\mathcal{Y}_a = 1.5 \cdot \left[\frac{B_{RF}}{B_b}\right]^3 \cdot \frac{1}{(2\pi k)^2} \cdot \mathcal{Y}_e \cdot \qquad (11)$$

Fig.10 shows $\mathcal{Y}_a$ as a function of $\mathcal{Y}_e$ with k as the parameter for the case when $B_{RF}/B_b = 10^3$ holds. It is apparent from Fig.10 that the jam resistance calculated from (11) decreases with increasing k. This is due to the fact that, in order to avoid degrading the operation of the tapped delay line correlator the bandwidth of the message modulation has to be restricted. A large k demands a small ß, see (10); however, a small ß causes a relatively small jam resistance, see (7).

Ideal synchronism is assumed in (11). Incorrect setting and the additional epoch error build-up between the setting pulses result in a further degradation. If τ_{rms} and the additional epoch error are kept well below $0.1 T_C$, this degradation is less than 1 dB.

3.2. SYSTEM IMPROVEMENT USING AN ADDITIONAL NARROW-BAND PHASE LOCKED LOOP

The out-of-lock time t_{out} of the system shown in Fig.5 could be reduced if the false alarm rate n_{fa} could be reduced without altering the detection probability P_d, see also (3). A reduction of n_{fa} in this manner is possible if the time intervals in which false alarms can occur are restricted. One possibility to achieve such a restriction is shown in Fig.11. In this figure the design shown in Fig.5 is supplemented by a narrow-band phase locked loop and a switch.

The phase locked loop synchronizes itself to the correlation peaks of the tapped delay line correlator. If the loop filter has a sufficiently narrow bandwidth, the function of the phase locked loop is virtually unaffected by the false alarms and losses at the output of the tapped delay line correlator. The output signal of the voltage controlled oscillator VCO allows the instants of the correlation peaks to be estimated with a time error which is much less than the period of the PN sequence. The output signal of the VCO can therefore be used to control a switch connected between the tapped delay line correlator and the local PN generator. The switch is closed during the time intervals within which the correlation peaks are expected. It is open for the rest of the time. This prevents the false alarms which occur outside a certain time interval centered about the estimated instant of the correlation peak from setting the local PN generator.

The interference susceptibility of the phase locked loop to false alarms and losses decreases as the cut-off frequency of the loop filter decreases. At the same time, however, the dynamics of the phase locked loop become poorer, and hence also the acquisition performance of the system as well as the time tolerance range for the closing of the switch. Corresponding optimization investigations are planned.

A further system improvement is possible by using two or more tapped delay line correlators which are matched to different sections of the PN sequence. Thereby the rate of the synchronization pulses can be increased. This might be desirable if long PN sequences are used.

4. PRACTICAL WORK

4.1. EXPERIMENTAL SET-UP

In order to demonstrate the operating capability of the system in Fig.5 an experimental communication system as shown in the block diagramm Fig.12 is presently being set up. The essential system data are given in Table I. The value m = 2 refers to an epoch error of $0.1 T_C$. Initial measurements have shown that m = 2 can be obtained with ease.

Transmission bandwidth	B_{RF} = 20 MHz	
Clock frequency of the PN sequences	f_C = 10 MHz	
Message bandwidth	B_b = 5 kHz	Table I. Data of the
Length of the PN sequence	L = 1023	experimental set-up
Length of the tapped delay line	k = 255	
Stability of the PN clock generators	m = 2	

4.2. DEMONSTRATION OF THE FEASIBILITY OF THE SYNCHRONIZATION METHOD WITH A BASEBAND CIRCUIT

The most essential components of the receiver are the tapped delay line correlator and local PN generator which together form the equipment for producing the synchronous local PN sequence. It is not necessary to set up the entire transmission system in order to demonstrate by experiment the operability of a synchronizing circuit of this type. It is simpler to do so with the aid of a baseband circuit as shown in Fig.13 in which only the PN sequence appears instead of a carrier modulated with the message and the PN sequence. The transmitter is modelled by a PN generator, the receiver by a CCD correlator and a settable PN generator. The CCD correlator, like the tapped delay line correlator of the original system, is matched to a section of the PN sequence. The correlation peaks occuring at the output of the CCD correlator are used to set and start the local PN generator. The received interfering signal in the original system can be simulated by a noise signal injected along with the PN sequence at the input of the CCD correlator.

The quality of the synchronism of the local PN sequence can be quantitatively determined by adding the local and the "transmitted" PN sequence modulo 2. If the two PN sequences are in perfect synchronism, the output signal of the mod-2-adder would have a constant logic value of L. If the local PN sequence is not synchronized exactly to the received PN sequence, the output signal of the modulo-2-adder jumps occasionally to the logic value H. The sum of the time intervals during which the output signal of the mod-2-adder has a

value of H be denoted by T_H.

The quotient

$$\frac{T_H}{T_O} = \frac{\text{overall duration of non-synchronism}}{\text{system operating time}}$$

is a measure of the quality of synchronism obtained. The authors have implemented a circuit as per Fig.13. The essential data of this circuit are given in Table II.

PN clock frequency (crystal)	$f_c = 2.6$ MHz	
Length of the PN sequences	$L = 1023$	Table II. Data of the
Number of taps of the CCD correlator	$k = 32$	set-up with CDD correlator
Amplitude of the PN sequence at input of the CCD correlator	± 1 V	

T_H/T_O was measured as a function of the signal-to-noise ratio γ at the input of the CCD correlator; the interfering signal was white noise with a cut-off frequency of 2.6 MHz. The threshold in the CCD correlator was so set that virtually no false alarms occured in the setting and starting signal with the signal-to-noise ratios under consideration. The T_H/T_O obtained from the measurements is plotted in Fig.14 as a function of γ. Also shown in this figure is the detection probability P_d of the setting and starting pulses as a function of γ. T_H/T_O is always in the region of a few percent. Sections of the disturbed PN sequence at the input of the CCD correlator as well as the PN sequences at the inputs of the mod-2-adder are shown in Fig.15 for the circuit as per Fig.13 and Table II for the case $\gamma = 1$. It is evident that the local PN sequence is not perfectly synchronous with the transmitted sequence, but is still close to it. The lag of the two lower runs in Fig. 15 with respect to the upper run is caused by signal delay in the test circuit.

APPENDIX I, DERIVATION OF EQUATIONS (3), (4) AND (5)

The probability of m successive synchronization pulses passing undetected is $(1 - P_d)^m$. In any such case the permissible epoch error is exceeded for a time interval of LT_C. $T_O/(LT_C)$ synchronization pulses occur during the entire transmission period T_O. Hence the detection losses contribute to t_{out} with the additive term $[T_O/(LT_C)](1-P_d)^m \cdot LT_C/T_O = (1-P_d)^m$. The time available during T_O for interfering false alarms is $T_O[1-(1-P_d)^m]$. If we assume $n_{fa}LT_C \ll 1$, synchronism is lost for an average period of $LT_C/2$ as a result of each false alarm. The false alarms therefore contribute to t_{out} with the additive term $[1-(1-P_d)^m] n_{fa}LT_C/2$. The two terms are added to produce (3).

During the synchronization pulses, the output signal $s_2(t)$ of the tapped delay line can be approximated to by a sinewave signal of amplitude kA, upon which a Gaussian interfering signal with power $A^2/(2\gamma_e k)$ is superimposed. The probability of the envelope curve of this signal exceeding a threshold value b is given by the Marcum-Q-function (STEIN, S. and JONES, J.J.,1967)

$$P_d = Q\left[\sqrt{2k\gamma_e}, \, b\sqrt{2\gamma_e}/(kA)\right].$$

By substituting the normalized threshold value $b' = b\sqrt{2\gamma_e}/kA$, we obtain expression (4). As long as there is no synchronization pulse at the output of the tapped delay line, output signal $s_2(t)$ can be approximated to by a bandpass noise signal $n(t) = n_d(t)\cos(\omega_o t) + n_q(t)\sin(\omega_o t)$. The signals $n_d(t)$ and $n_q(t)$ are assumed to be uncorrelated Gaussian processes with an average value of zero which are limited to the frequency range $|f| \leq f_c$ and which over this range have a two sided power density of $S_n(f) = kA^2(1+1/\gamma_e)/(4f_c)$. $S_n(f)$ contains a component arising from the received interfering signal and another arising from the received desired signal. The square of the envelope of $n(t)$ has the p.d.f.

$$p(v) = \frac{1}{kA^2(1+1/\gamma_e)} \cdot \exp\left(\frac{-v}{kA^2(1+1/\gamma_e)}\right).$$

Let us assume that the autocorrelation function of $n_d(t)$ or $n_q(t)$ is $R(\tau)$. If we differentiate $R(\tau)$ twice with respect to τ and substitute $\tau = 0$, we obtain

$$R''(0) = -4\pi^2 \int_{-f_c}^{f_c} f^2 S_n \, df = -\frac{2}{3}\pi^2 f_c^2 \, kA^2(1+1/\gamma_e).$$

The rate at which the envelope curve of $n(t)$ exceeds a threshold value b in accordance with (HELSTROM, C.W.,1968) with $p(v)$ and $R''(0)$, is

$$n_{fa} = p(b^2)\sqrt{2|R''(0)|} \, b^2/\pi = \sqrt{\frac{2\pi}{3}} \cdot \frac{b \, f_c}{\sqrt{kA^2(1+1/\gamma_e)/2}} \exp\left[\frac{-b^2}{kA^2(1+1/\gamma_e)}\right].$$

Substituting in this equation $f_c = 1/T_C$ and the normalized threshold value $b' = b\sqrt{2\gamma_e}/(kA)$ we obtain equation (5).

APPENDIX II, DERIVATION OF EQUATION (6)

The rectifier of the tapped delay line correlator is assumed to be an envelope detector. In this case the desired component of output signal $s_3(t)$ can be approximated to as shown in Fig.7. The envelope in Fig.7 has a maximum value of kA. If there is also an interfering signal at the input of the tapped delay line, an interfering signal is superimposed on output signal $s_3(t)$. Its power is approximately given by

$$\sigma^2 = kA^2/(2\gamma_e) \ ,$$

for sufficiently large $k\gamma_e$, see (STEIN, S. and JONES, J.J.,1967). The edge slope of the desired signal as per Fig. 7 is $d = kA(T_c/2) = 2kA/T_c$. According to (HÖLZLER, E. and HOLZWARTH, H.,1975) the variation of the instant when the threshold is exceeded is then

$$\tau^2_{rms} = \frac{\sigma^2}{d^2} = \frac{kA^2/(2\gamma_e)}{(2kA/T_c)^2} = \frac{T_c^2}{8k\gamma_e} \ .$$

Equation (6) is obtained from this by dividing by T_c and taking the square root.

APPENDIX III, DERIVATION OF EQUATION (7)

Since frequency modulation with modulation index β is used for message modulation, the bandwidth of the interference suppressing bandpass filter in the receiver must be made $B'_{RF} = 2 \cdot B_b \cdot (1+\beta)$ (STEIN, S. and JONES, J.J.,1967). Accordingly, the processing gain (CAHN, C.R. et al.,1973; DIXON, R.C.,1976) is $B_{RF}/B'_{RF} = B_{RF}/[2B_b(1+\beta)]$. The signal-to-noise ratio is further improved by a factor of $3\beta^2(1+\beta)$ by the frequency demodulation (STEIN, S. and JONES, J.J.,1967). The signal-to-noise ratio γ_a at the demodulator output is hence obtained from the signal-to-noise ratio γ_e at the receiver input with the aid of the relationship $\gamma_a = B_{RF} \cdot 3\beta^2(1+\beta)\gamma_e/[2B_b(1+\beta)]$. This is equation (7).

REFERENCES

BAIER, W.P., 1975, "Überlegungen zu störsicheren drahtlosen Nachrichtenübertragungssystemen", Siemens Forsch.- und Entwickl.- Ber. Bd. 4 (1975), 61-67.

BAIER, W.P. and MEFFERT, K., 1977, "Optimale Taktfrequenz bei Nachrichtenübertragungssystemen mit pseudozufälliger Phasensprungmodulation", Frequenz 31(1977)8, 243-246.

BELL, D.T. et al., 1973, "Application of Acoustic Surface-Wave Technology to Spread Spectrum Communications", IEEE Trans. Microwave Theory Techn., vol. MTT-21 (1973), 263-271.

BUSS, B.D. et al., 1973, "Transversal Filtering Using Charge-Transfer Devices", IEEE Journal Solid-State Circuits, vol. SC-8 (1973), 138-146.

CAHN, C.R. et al., 1973, "Spread Spectrum Communications", AGARD LSP-58.

DAVIES, N.G. and WOHLBERG, D.B., 1976, "Spread Spectrum Techniques", Proceedings of the International Specialist Seminar on the Impact of New Technologies in Signal Processing, Aviemore, Scotland, IEE Conference Publication no. 144, 144-154.

DIXON, R.C., 1976, "Spread Spectrum Systems", New York, John Wiley & Sons.

GILL, W.J., 1966, "A Comparison of Binary Delay-Lock Tracking-Loop Implementations", IEEE Trans Aerosp. Electr. Sys., vol. AES-2 (1966), 415-424.

HAYS, R.M. and HARTMANN, C.S., 1976, "Surface-Acoustic-Wave Devices for Communications", Proceedings of the IEEE, vol. 64 (1976), 652-671.

HELSTROM, C.W., 1968, "Statistical Theory of Signal Detection", Pergamon Press.

HÖLZLER, E. and HOLZWARTH, H., 1975, "Pulstechnik" vol. 1, Springer Verlag.

SAGE, G.F., 1964, "Serial Synchronization of Pseudonoise Systems", IEEE Trans. Commun. Technol., vol. COM-12 (1964), 123-127.

STEIN, S. and JONES, J.J., 1967, "Modern Communication Principles", Mc Graw Hill.

WARD, R.B., 1965, "Acquisition of Pseudonoise Signals by Sequential Estimation", IEEE Trans. Commun. Technol. vol. COM-13 (1965), 475-483.
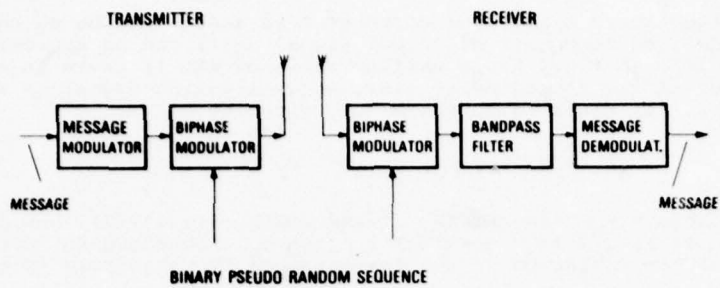
Fig. 1. Direct sequencing communication system
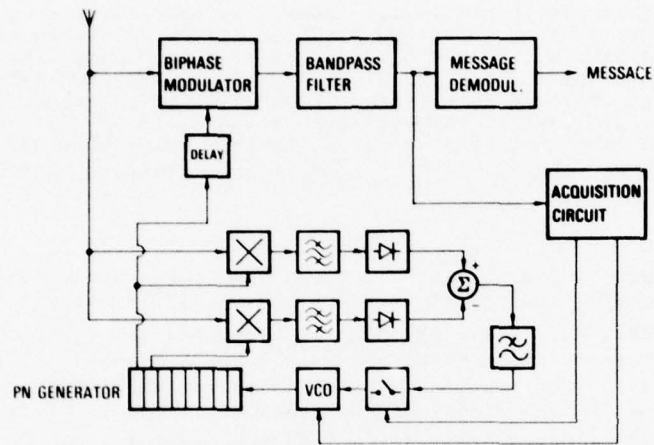
Fig. 2. Acquisition by active correlation
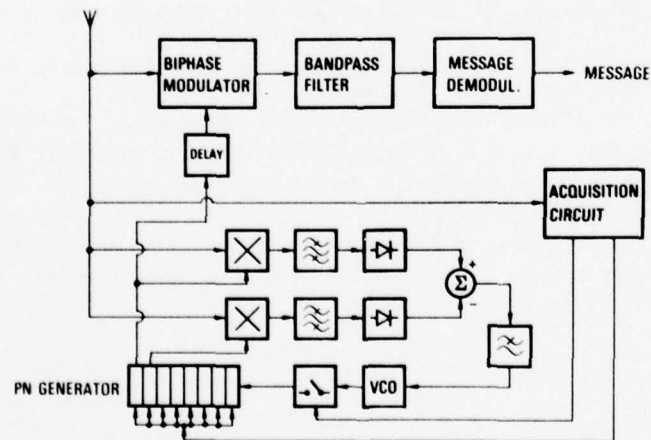
Fig. 3. Acquisition by passive correlation
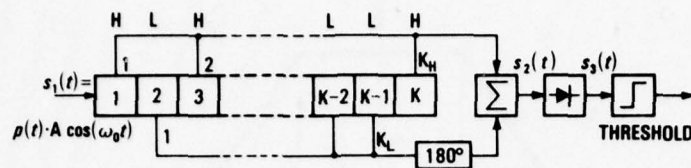
Fig. 4. Tapped delay line correlator
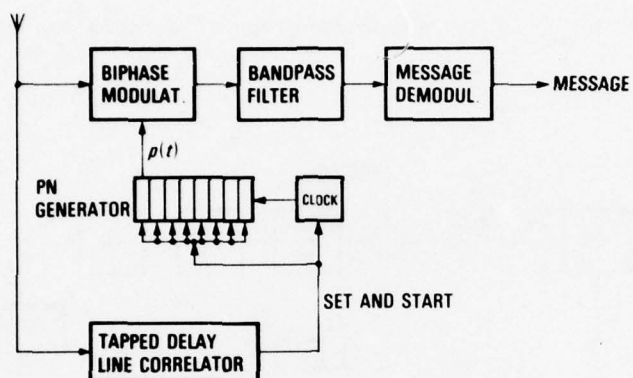
Fig. 5. Combined acquisition and tracking circuit
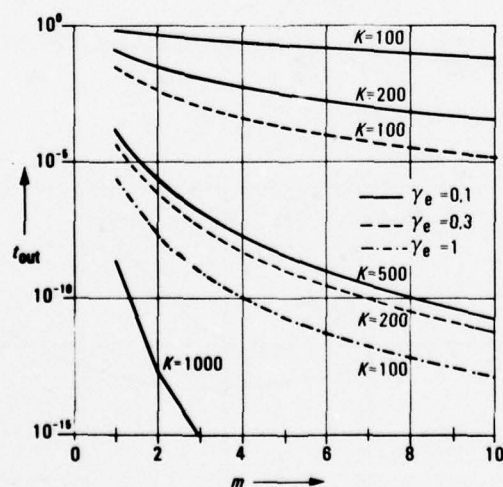
Fig. 6. Out-of-Lock t_{out} time t_{out} versus m
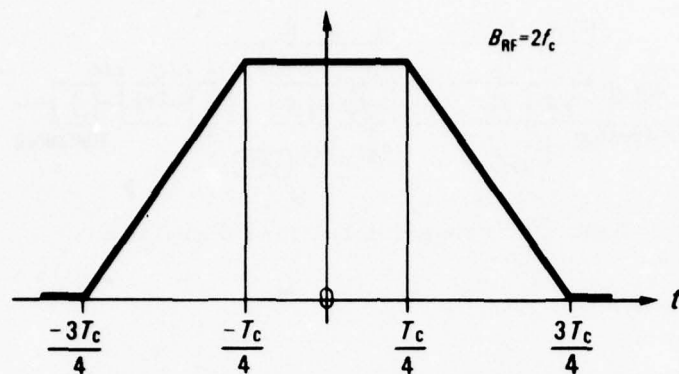
Fig. 7. Approximate envelope of correlation peak
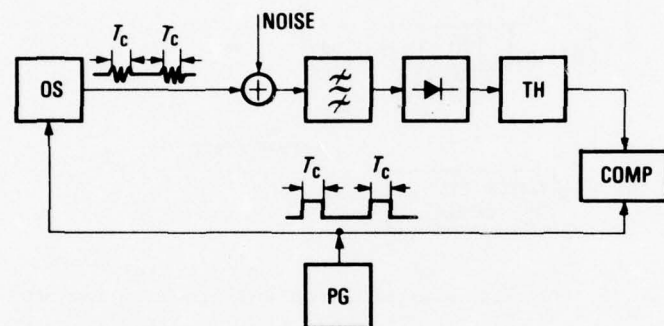
PG: PULSE GENERATOR
OS: OSCILLATOR
$\approx$: BANDPASS FILTER, 3dB-BANDWIDTH $2/T_c$
TH: ADJUSTABLE THRESHOLD CIRCUIT
COMP: CIRCUIT FOR MEASURING τ_{rms}

Fig. 8. Circuit for measuring the standard deviation of time accuracy
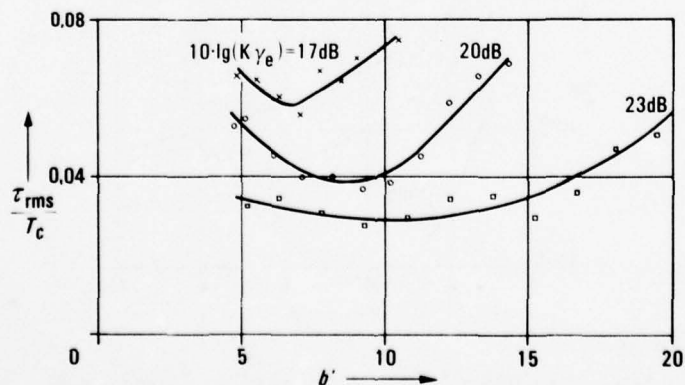
Fig. 9. Measured standard deviation of time accuracy
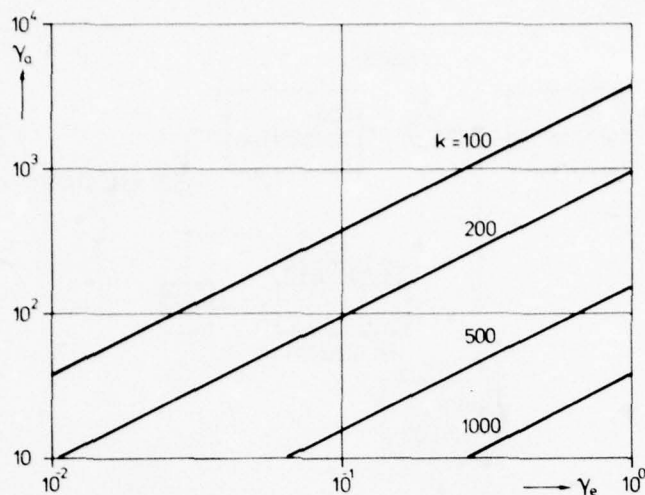
Fig. 10. Message signal-to-noise-ratio γ_a versus signal-to-noise-ratio γ_e at the receiver input
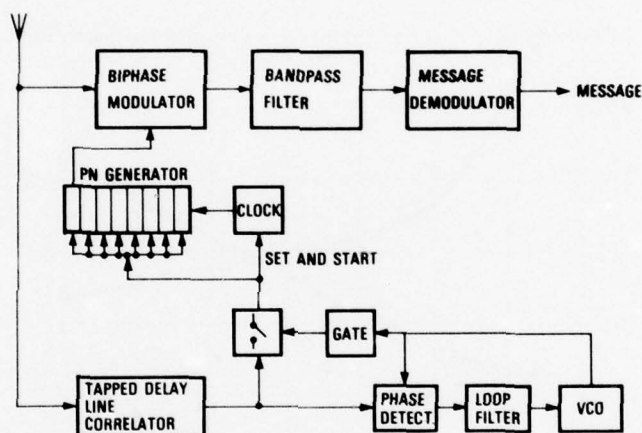
Fig. 11. Improved combined acquisition and tracking circuit
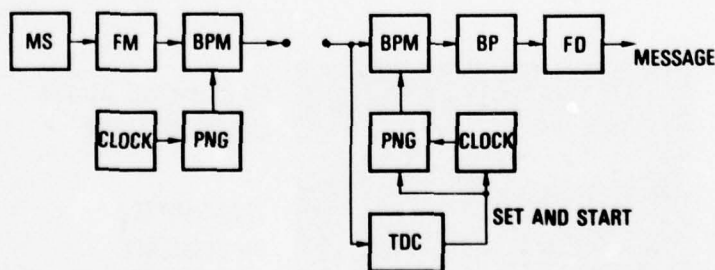
MS: MESSAGE SOURCE (0...5kHz)

FM: FREQUENCY MODULATOR
($\beta=0...10$, $f_0\approx75\,MHz$)

PNG: PN GENERATOR
($L=1023$, $f_c\approx10\,MHz$)

BPM: BIPHASE MODULATOR

BP: BANDPASS FILTER ($B=100kHz$)

FD: FREQUENCY DEMODULATOR

TDC: TAPPED DELAY LINE CORRELATOR

Fig. 12. Experimental set-up with SAW tapped delay line correlator
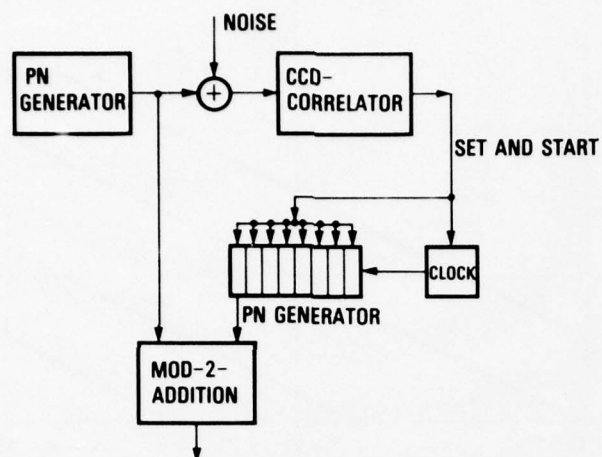
Fig. 13. Experimental set-up with CCD tapped delay line
correlator
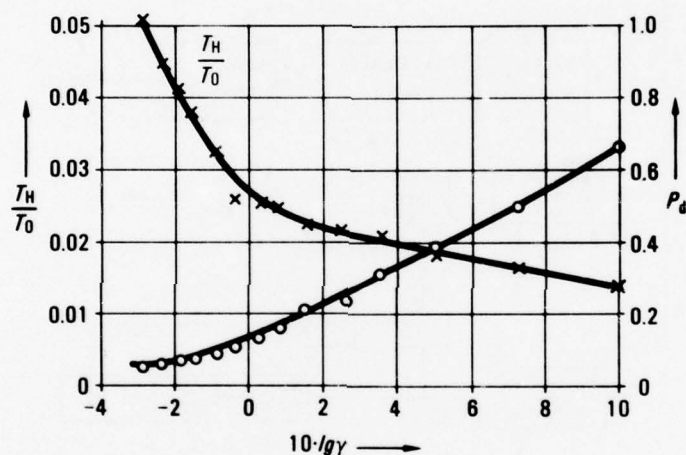
Fig. 14. T_H/T_O and P_d versus signal-to-noise ratio for the
synchronization circuit with CCD correlator

PN SEQUENCE AT THE
CCD CORRELATOR

TRANSMITTED
PN SEQUENCE

LOCAL PN SEQUENCE
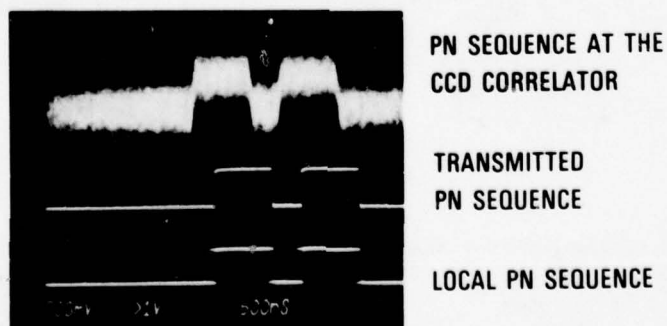
Fig. 15. PN sequences in the synchronization circuit with
CCD correlator

AD-A058 838   ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT--ETC F/G 9/5
IMPACT OF CHARGE COUPLED DEVICES AND SURFACE ACOUSTIC WAVE DEVI--ETC(U)
OCT 77   Y BRAULT

UNCLASSIFIED   AGARD-CP-230   NL

6 OF 6
AD
A058 838

END
DATE
FILMED
-11 -78
DDC

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document
\	AGARD-CP-230	ISBN 92-835-0216-7	UNCLASSIFIED

5. Originator	Advisory Group for Aerospace Research and Development North Atlantic Treaty Organization 7 rue Ancelle, 92200 Neuilly sur Seine, France

6. Title	IMPACT OF CHARGE COUPLED DEVICES AND SURFACE ACOUSTIC WAVE DEVICES ON SIGNAL PROCESSING AND IMAGERY IN ADVANCED SYSTEMS

7. Presented at	the Avionic Panel Symposium held in Ottawa, Canada, 11–15 October 1977.

8. Author(s)		9. Date
Various	Edited by Y.Brault*	June 1977

10. Author's Address	*Thomson CSF Division Equipements Avioniques & Spatiaux 178 Bld Gabriel Péri, 92240 Malakoff	11. Pages
Various		502

12. Distribution Statement	This document is distributed in accordance with AGARD policies and regulations, which are outlined on the Outside Back Covers of all AGARD publications.

13. Keywords/Descriptors		
Charge coupled devices	Images	Avionics
Surface acoustic waves	Digital systems	Meetings
Signal processing	Display devices	

14. Abstract

New concepts in imagery and signal processing techniques have been fostered by the technological advances in charge coupled devices and surface acoustic waves. These two technologies have not, in general, been treated in a common seminar.

This Symposium was an attempt to promote in each community an awareness of the concepts and techniques which could lead to a synthesis of the two in advanced digital systems.

AGARD Conference Proceedings No.230
Advisory Group for Aerospace Research and Development, NATO
IMPACT OF CHARGE COUPLED DEVICES AND SURFACE ACOUSTIC WAVE DEVICES ON SIGNAL PROCESSING AND IMAGERY IN ADVANCED SYSTEMS
Edited by Y.Brault
Published June 1978
502 pages

New concepts in imagery and signal processing techniques have been fostered by the technological advances in charge coupled devices and surface acoustic waves. These two technologies have not, in general, been treated in a common seminar.

P.T.O.